THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 — NUMBER 1

CONTENTS

JANUARY 1966

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

THIS IS THE OFFICIAL JOURNAL OF THE

MATHEMATICAL ASSOCIATION OF AMERICA, INC.

(Devoted to the Interests of Collegiate Mathematics)

General Office: SUNY at Buffalo, Buffalo, N. Y. 14214. *Executive Director:* H. M. GEHMAN

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA
(INCORPORATED)

DEVOTED TO THE INTERESTS OF COLLEGIATE MATHEMATICS

THE AMERICAN MATHEMATICAL MONTHLY, FOUNDED IN 1894 BY BENJAMIN F. FINKEL, WAS PUBLISHED BY HIM UNTIL 1913. FROM 1913 TO 1916 IT WAS OWNED AND PUBLISHED BY REPRESENTATIVES OF FOURTEEN UNIVERSITIES AND COLLEGES IN THE MIDDLE WEST

VOLUME 73
1966

PUBLISHED BY THE ASSOCIATION
MENASHA, WIS., AND BUFFALO, N. Y.

AWARD FOR DISTINGUISHED SERVICE TO PROFESSOR HARRY MERRILL GEHMAN

Fittingly at this time the Mathematical Association of America honors Harry Merrill Gehman by naming him recipient of the Award for Distinguished Service to Mathematics. The dynamic strength of our Association as it now enters its second half-century can be attributed in marked measure to his devoted and diligent service—for twelve years (1948–60) as its secretary-treasurer and, since 1960, as its treasurer and first executive director. With prudence and patience, good sense and good humor, he has guided the diverse activities of the Association through a difficult period of burgeoning growth. Today we are veritably saluting Mister MAA!

Born at Norristown, Pennsylvania, in 1898, Harry Gehman did his undergraduate and graduate work at the University of Pennsylvania and began his teaching there. His A.B. (with ΦBK) in 1919 and A.M. in 1920 came just at the time Professors R. L. Moore and J. R. Kline had made Pennsylvania the first home of American point-set topology (then called Analysis Situs). While serving as an instructor at Pennsylvania, Harry Gehman took a hand in this exciting new development and obtained his Ph.D. in 1925 with a thesis written under the supervision of J. R. Kline. After a year as National Research Council fellow with R. L. Moore at the University of Texas, he taught for three years at Yale University, one as instructor and two as assistant professor.

In 1929 he accepted a professorship at the University of Buffalo, where Chancellor Capen had attracted promising scholars in several fields to establish a tutorial plan in the University's new College of Arts and Sciences. He threw himself vigorously into building a strong mathematical program at Buffalo, while at the same time continuing active research in the topology of continuous curves [see bibliography in R. L. Moore, *Foundations of Point Set Theory*, American Mathematical Society Colloquium Publication No. 13]. For thirty-three years (1929–1962) he served as head of the Department of Mathematics at the University of Buffalo, now the State University of New York at Buffalo. There was one interlude: in 1945, literally at a moment's notice, he went to England to head the Department of Mathematics of the Shrivenham University, United States Army.

Harry Gehman was the first chairman (1940–41) of the Upper New York State section of the Association—an experience which seems to have led him to strive constantly to make the regional sections of the Association the "grass roots" of its nation-wide strength. He was elected to the MAA Finance Committee in 1944, and succeeded W. B. Carver as secretary-treasurer in 1948. By effective selfless service, in the best tradition of his predecessors, he became the mainstay of our Association—managing its financial matters, implementing decisions of the Board of Governors, coordinating committee activities, and supervising closely both editing and publishing.

HARRY MERRILL GEHMAN

In addition to serving ex officio on the Board of Governors and various Association committees (the Executive and Finance Committees, the Committees on Publications and on Membership, and the Joint Committee on Places of Meetings), he has served by appointment on nominating committees, on the planning committee for the fiftieth anniversary celebration, and as chairman of the Committee on the Structure of the Government of the Association. He has represented the Association in the Division of Mathematics of the National Research Council, on the Policy Committee for Mathematics and then the Conference Board of the Mathematical Sciences, where he was elected to an additional term as member-at-large. He was Chairman in 1951–52 of the mathematics section of the American Society for Engineering Education and participated three summers (1959–61) in writing projects of the School Mathematics Study Group.

Over and beyond this handsome record of accomplishment shine the amiable warmth and sterling spirit of the man himself—and his wife, Marian, who has shared in his mission to mathematics. It has been service with the extra zeal and zest of a man who enjoys working with other people and accepts their foibles with a chuckle. For all this we proudly and gratefully hail Harry Merrill Gehman.

A. W. TUCKER

AVOIDING THE JORDAN CANONICAL FORM IN THE DISCUSSION OF LINEAR SYSTEMS WITH CONSTANT COEFFICIENTS

E. J. PUTZER, North American Aviation Science Center, Thousand Oaks, California
(Now at 5143 Topanga Canyon, Woodland Hills, California)

Consider the differential equation

$$\dot{x} = Ax; \qquad x(0) = x_0; \qquad 0 \leqq t < \infty, \tag{1}$$

where x and x_0 are n-vectors and A is an $n \times n$ matrix of constants. In this paper we present two methods, believed to be new, for explicitly writing down the solution of (1) without making any preliminary transformations. This is particularly useful, both for teaching and applied work, when the matrix A cannot be diagonalized, since the necessity of discussing or finding the Jordan Canonical Form (J.C.F.) of A is completely by-passed.

If e^{At} is defined as usual by a power series it is well known (see [1]) that the solution of (1) is

$$x = e^{At}x_0,$$

so the problem is to calculate the function e^{At}. In [2], this is done via the J.C.F. of A. In [1], it is shown how the J.C.F. can be by-passed by a transformation which reduces A to a triangular form in which the off diagonal elements are arbitrarily small. While this approach permits a theoretical discussion of the form of exp $\{At\}$ and its behavior as t becomes infinite (Theorem 7 of [1]), it is not intended as a practical method for calculating the function. The following two theorems suggest an alternate approach which can be used both for calculation and for expository discussion. It may be noted that the formula of Theorem 2 is simpler than that of Theorem 1 since the r_i are easier to calculate than the q_i.

In order to state Theorem 1 simply, it will be convenient to introduce some notation.

Let A be an $n\times n$ matrix of constants, and let

$$f(\lambda) \equiv |\lambda I - A| \equiv \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

be the characteristic polynomial of A. Construct the scalar function $z(t)$ which is the solution of the differential equation

$$z^{(n)} + c_{n-1}z^{(n-1)} + \cdots + c_1\dot{z} + c_0 z = 0 \tag{2}$$

with initial conditions

$$z(0) = \dot{z}(0) = \cdots = z^{(n-2)}(0) = 0; \qquad z^{(n-1)}(0) = 1. \tag{3}$$

We observe at this point that regardless of the multiplicities of the roots of $f(\lambda) = 0$, once these are obtained it is trivial to write down the general solution of (2). Then one solves a single set of linear algebraic equations to satisfy the initial conditions (3). Since the right member of each of these equations is zero except for the last, solving them entails only finding the cofactors of the elements of the *last row* of the associated matrix. It is not necessary to invert the matrix itself. For teaching purposes, the point is that the form of the general solution of (2) can be obtained quickly and easily by elementary methods.

Now define

$$Z(t) = \begin{bmatrix} z(t) \\ \dot{z}(t) \\ \cdot \\ \cdot \\ z^{(n-1)}(t) \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} c_1 & c_2 & \cdots & c_{n-1} & 1 \\ c_2 & c_3 & \cdots & 1 & \\ \cdot & & \cdot & & \\ \cdot & \cdot & & & 0 \\ c_{n-1} & 1 & & & \\ 1 & & & & \end{bmatrix}$$

We then have the following

THEOREM 1.

$$e^{At} = \sum_{j=0}^{n-1} q_j(t) A^j, \tag{4}$$

where $q_0(t), \cdots, q_{n-1}(t)$ *are the elements of the column vector*

$$\mathfrak{q}(t) = CZ(t). \tag{5}$$

Before we prove this, a remark is in order as to what happens when $f(\lambda)$ has multiple roots but the minimal polynomial of A has distinct factors so that A can in fact be diagonalized. It appears at first glance that our formula will contain powers of t, yet we know this cannot be the case. What occurs, of course, is that the powers of t in (4) just cancel each other out. This is the nice feature of formula (4); it is true for *all* matrices A, and we never have to concern ourselves about the nature of the minimal polynomial of A, or its J.C.F., and no preliminary transformations of any kind need be made.

Proof. We will show that if

$$\Phi(t) = \sum_{j=0}^{n-1} q_j(t) A^j \tag{6}$$

then $d\Phi/dt = A\Phi$ and $\Phi(0) = I$, so $\Phi(t) = e^{At}$. Since only $q_0(t)$ involves $z^{(n-1)}(t)$, $q_j(0) = 0$ for $j \geqq 1$. Clearly, $q_0(0) = 1$. Thus, $\Phi(0) = I$.

Now consider $(d\Phi/dt) - A\Phi$. Differentiating (6), and applying the Hamilton-Cayley Theorem

$$A^n + \sum_{j=0}^{n-1} c_j A^j = 0,$$

we obtain

$$\frac{d\Phi}{dt} - A\Phi = (\dot{q}_0 + c_0 q_{n-1}) + \sum_{j=0}^{n-1} (\dot{q}_j - q_{j-1} + c_j q_{n-1}) A^j.$$

It will suffice, therefore, to show that

$$\begin{aligned} \dot{q}_0(t) &\equiv -c_0 q_{n-1}(t), \\ \dot{q}_j(t) &\equiv q_{j-1}(t) - c_j q_{n-1}(t) \qquad j = 1, \cdots, (n-1). \end{aligned}$$

From the definition (5),

$$q_j(t) \equiv \sum_{k=1}^{n-j-1} c_{k+j} z^{(k-1)} + z^{(n-j-1)}. \tag{7}$$

Therefore $\dot{q}_j(t) \equiv \sum_{k=1}^{n-j-1} c_{k+j} z^{(k)} + z^{(n-j)}$.

But $q_{n-1} \equiv z$, so we have

$$\dot{q}_j + c_j q_{n-j} \equiv \sum_{k=0}^{n-j-1} c_{k+j} z^{(k)} + z^{(n-j)} \quad \text{for } j = 0, 1, \cdots, n-1. \tag{8}$$

If $j = 0$, this yields

$$\dot{q}_0 + c_0 q_{n-1} \equiv \sum_{k=0}^{n-1} c_k z^{(k)} + z^{(n)}$$

which is zero because of (2).

If $j \geqq 1$, replace j by $j-1$ in (7) and change the summation index from k to $k+1$ to get

$$q_{j-1}(t) \equiv \sum_{k=0}^{n-j-1} c_{k+j} z^{(k)} + z^{(n-j)}. \tag{9}$$

Comparing (9) and (8) we have $\dot{q}_j + c_j q_{n-1}(t) \equiv q_{j-1}(t)$ for $j=1, 2, \cdots, n-1$.

Students will want to see the formula (4) *derived.* One merely presents the proof backward, beginning with the observation that because of the Hamilton-Cayley theorem, exp $\{At\}$ should be expressible in the form (6). Regarding the q_j as unknowns, and applying the differential equation which exp $\{At\}$ satisfies, leads directly to (4).

A second explicit formula for exp $\{At\}$, which also holds for all matrices A, is the following: Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the eigenvalues of A in some arbitrary but specified order. These are not necessarily distinct. Then

THEOREM 2. $e^{At} = \sum_{j=0}^{n-1} r_{j+1}(t) P_j$, *where*

$$P_0 = I; \qquad P_j = \prod_{k=1}^{j} (A - \lambda_k I), \qquad (j = 1, \cdots, n),$$

and $r_1(t), \cdots, r_n(t)$ *is the solution of the triangular system*

$$\begin{cases} \dot{r}_1 = \lambda_1 r_1 & \\ \dot{r}_j = r_{j-1} + \lambda_j r_j & (j = 2, \cdots n) \\ r_1(0) = 1; \quad r_j(0) = 0 & (j = 2, \cdots n). \end{cases}$$

Proof. Let

$$\Phi(t) \equiv \sum_{j=0}^{n-1} r_{j+1}(t) P_j \tag{10}$$

and define $r_0(t) \equiv 0$. Then from (10) and the equations satisfied by the $r_j(t)$ we have, after collecting terms in r_j,

$$\dot{\Phi} - \lambda_n \Phi = \sum_{0}^{n-2} [P_{j+1} + (\lambda_{j+1} - \lambda_n) P_j] r_{j+1}.$$

Using $P_{j+1} \equiv (A - \lambda_{j+1} I) P_j$ in this gives

$$\begin{aligned} \dot{\Phi} - \lambda_n \Phi &= (A - \lambda_n I)(\Phi - r_n(t) P_{n-1}) \\ &= (A - \lambda_n I)\Phi - r_n(t) P_n. \end{aligned}$$

But $P_n \equiv 0$ from the Hamilton-Cayley Theorem, so $\dot{\Phi} = A\Phi$. Then since $\Phi(0) = I$ it follows that $\Phi(t) = e^{At}$.

Example. If one desires a numerical example for class presentation an ap-

propriate matrix A can be prepared in advance by arbitrarily choosing a set of eigenvalues, a Jordan Canonical Form J and a nonsingular matrix S and calculating

$$A = SJS^{-1}.$$

Then beginning with A, one simply calculates the set $\{q_i(t)\}$ and/or $\{r_i(t)\}$. Consider the case of a 3×3 matrix having eigenvalues (λ, λ, μ). There are two subcases; the one in which the normal form of A is diagonal, and the one in which it is not. These two subcases are taken care of automatically by the given formula for exp $\{At\}$, and do not enter at all into the calculation of the $\{q_i\}$ or $\{r_i\}$.

As an example we will explicitly find the sets $\{q_i\}$ and $\{r_i\}$ for the case of a 3×3 matrix with eigenvalues $(\lambda, \lambda, \lambda)$. We note that aside from the trivial case in which the normal form (and hence A itself) is diagonal, there are two distinct nondiagonal normal forms that A may have. As above, these do not have to be treated separately.

From Theorem 1,

$$f(x) \equiv (x-\lambda)^3 = x^3 - 3\lambda x^2 + 3\lambda^2 x - \lambda^3$$

so $c_1 = 3\lambda^2$, $c_2 = -3\lambda$. Obviously $z(t) = (a_1 + a_2 t + a_3 t^2)e^{\lambda t}$. Applying the initial conditions to find the a_i yields $z = \frac{1}{2}t^2 e^{\lambda t}$, so

$$Z(t) = \tfrac{1}{2}e^{\lambda t}\begin{bmatrix} t^2 \\ \lambda t^2 + 2t \\ \lambda^2 t^2 + 4\lambda t + 2 \end{bmatrix}.$$

Then since

$$C = \begin{bmatrix} 3\lambda^2 & -3\lambda & 1 \\ -3\lambda & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$q = CZ(t) = \tfrac{1}{2}e^{\lambda t}\begin{bmatrix} \lambda^2 t^2 - 2\lambda t + 2 \\ -2\lambda t^2 + 2t \\ t^2 \end{bmatrix}.$$

Thus

$$e^{At} = \tfrac{1}{2}e^{\lambda t}\{(\lambda^2 t^2 - 2\lambda t + 2)I + (-2\lambda t^2 + 2t)A + t^2 A^2\} \tag{11}$$

for every 3×3 matrix A having all three eigenvalues equal to λ. The corresponding formula from Theorem 2 is obtained by solving the system

$$\begin{cases} \dot r_1 = \lambda r_1 \\ \dot r_2 = r_1 + \lambda r_2 \\ \dot r_3 = r_2 + \lambda r_3 \end{cases}$$

with the specified initial values. This immediately gives

$$r_1 = e^{\lambda t}; \qquad r_2 = te^{\lambda t}; \qquad r_3 = \frac{t^2}{2} e^{\lambda t}$$

so

$$e^{At} = \tfrac{1}{2}e^{\lambda t}\{2I + 2t(A - \lambda I) + t^2(A - \lambda I)^2\}. \tag{12}$$

Of course, if we collect powers of A in (12) we will obtain (11).

References

1. R. E. Bellman, Stability Theory of Differential Equations, McGraw-Hill, New York, 1953
2. E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

THE IMPORTANCE OF ASYMPTOTIC ANALYSIS IN APPLIED MATHEMATICS

L. A. SEGEL, Rensselaer Polytechnic Institute

Friedrichs began his 1954 Gibbs lecture on asymptotic phenomena in mathematical physics [6] by saying "the problems I intend to speak about belong to the somewhat undefined and disputed region at the border between mathematics and physics." In the ten years since his lecture there has been a growing tendency to define this region and some others as the domain of applied mathematics. A number of explicit university applied mathematics programs have been started, the Society for Industrial and Applied Mathematics has been founded in the United States and the Institute of Mathematics and its Applications in England.

It ought to be helpful to have on record various opinions on what this new, or renewed, activity is all about. The general nature of applied mathematics has been ably discussed [7], [8] and the time now appears ripe for more detailed remarks on the elements of an applied mathematical outlook. The purpose of the following supplement to Friedrichs' remarks and references is to stress the importance of asymptotic analysis in applied mathematics. Although they will not be mentioned here, there are of course several other important aspects of the subject. One of these is the use of computing machines. Murray [13] concentrates so exclusively on this one aspect that he gives the impression that applied mathematics is a problem solving service rather than the independent science envisioned here and elsewhere. An indication of the central position of asymptotics is that in concentrating on it we do not present an unbalanced picture of applied mathematics.

Another important aspect of the subject is regular perturbation theory, which gives rise to convergent series solutions to problems. A convergent power series about a point is asymptotic as the independent variable approaches the point, but we wish to show the importance of "genuine" or nonconvergent asymptotic series. To do this we shall briefly examine (i) certain basic theorems connected with asymptotic expansions, (ii) a differential equation in the neighborhood of an essential singularity, (iii) a very simple singular perturbation problem, and (iv) the lack of genuine distinction between a large variable and a small one—all in order to emphasize the connection between asymptotic approximations and essential singularities. By means of some examples, we shall try to show that the principal goal of the applied mathematician, basic understanding of physical phenomena, is most likely to be attained when asymptotic analysis is used to obtain the qualitative behavior of a solution near its worst (essential) singularity.

In his lecture, Friedrichs discussed two permissible views of Stokes-phenomenon-like discontinuities—the genuine boundary of a natural object or an approximation to a region of rapid change. He mentioned the paradox that classical mechanics is the first asymptotic approximation to quantum mechanics yet the latter cannot be defined without reference to the former. He thereby gave physical support for the view that asymptotic methods offer more than another approximation technique but rather have a fundamental role in the mathematical description of physical phenomena. This same view is supported here by stressing the connection between asymptotics and essential singularities.

By definition, $f(z)$ has an asymptotic power series P near z_0,

$$f(z) \sim P, \qquad P \equiv \sum_{n=0}^{\infty} A_n(z - z_0)^n,$$

if for every N

$$\left[f(z) - \sum_{n=0}^{N} A_n(z - z_0)^n\right] = O(z - z_0)^{N+1} \quad \text{as} \quad z \to z_0$$

or

$$\lim_{z \to z_0} (z - z_0)^{-N}\left[f(z) - \sum_{n=0}^{N} A_n(z - z_0)^n\right] = 0. \tag{1}$$

(Most often z_0 is the point at infinity and the definition is modified in an obvious way. See [3], a basic reference for asymptotics.) Also $f(z) \sim g(z)P$ if and only if $f/g \sim P$. More general asymptotic series have sometimes proved useful, particularly in discussing functions of two complex variables [5].

Although different functions can have the same asymptotic power series, it follows from (1) that if a given function has a series the series is unique:

(2a) $A_0 = \lim_{z \to z_0} f(z)$, (2b) $A_1 = \lim_{z \to z_0} (z - z_0)^{-1}[f(z_0) - A_0]$, etc.

In practice it is frequently clear that f has an isolated essential singularity at z_0, so one might feel that (2a) conflicts with Weierstrass' theorem that $f(z)$ tends to any limit whatever as $z \to z_0$ through an appropriate sequence of values [**15**, p. 93]. Usually, however, we wish to approach z_0 along a continuous path, not along a discrete sequence of points. To see what now happens, let z_0 be the only singularity of the (entire) function f; f is defined to have *order* ρ if ρ is the greatest lower bound of positive numbers A such that

$$f(z) = O\left(\exp |z - z_0|^A\right), \quad z \to z_0.$$

It can be shown that an entire function of finite order ρ approaches at most 2ρ different values if $z \to z_0$ along some continuous curve [**15**, pp. 284–6]. This theorem lies behind the fact that different asymptotic series are usually encountered in different sectors of the complex plane (Stokes phenomenon). Together with the result that if f is analytic and single-valued for z near z_0, and if (1) holds for all arg $(z - z_0)$, then P converges to f near z_0 [**3**, p. 22], it shows that asymptotic series are to be anticipated when approximating a function in the neighborhood of an essential singularity.

"Asymptotic approximations are a type of approximation to a function that holds when a variable is large" writes Jeffries in the first sentence of his valuable book on asymptotics [**9**]. Essentially the same thought begins Rosser's survey article on the subject [**14**]. Presumably for simplicity in exposition, these authors choose to ignore examples like

$$\text{(3)} \qquad \epsilon \frac{d^2y}{dt^2} + 2\frac{dy}{dt} + y = 0, \qquad y(0) = 0, \qquad y(1) = 1,$$

or

$$\text{(4)} \qquad t^4 \frac{d^2y}{dt^2} + t^3 \frac{dy}{dt} + (1 - p^2t^2)y = 0, \qquad p \text{ a constant},$$

where asymptotic approximations are associated with small values of the parameter ϵ or the independent variable t. We expect, correctly, that expansions of solutions to (4) about the origin will be asymptotic because $t = 0$ is an irregular singular point of the differential equation so the general solution has an essential singularity at $t = 0$. (On p. 533 of [**12**], Morse and Feshbach have a suggestive discussion of the relation between singular points of the equation and its solutions.) After the transformation $x = 1/t$, (4) is the familiar Bessel equation

$$\text{(5)} \qquad x^2Y'' + xY' + (x^2 - p^2)Y = 0, \qquad Y(x) \equiv y(1/x).$$

Many students meet asymptotic series for the first time in advanced calculus treatments of (5) for large x. Unfortunately the impression is often left that a mysterious nonconvergence of the series is connected with proximity to the mysterious point at infinity.

If ϵ is small in (3), the natural thing to do is to ignore the term $\epsilon y''$. A first order equation results, making it impossible to satisfy both boundary conditions. This is one of the simplest examples of a singular perturbation problem in which the equation of the natural first approximation is qualitatively different from the full equation. That asymptotic approximations are associated with singular perturbation problems is suggested by the exact solution to (3):

$$y(t, \epsilon) = \frac{\exp(m_1 t) - \exp(m_2 t)}{\exp(m_1) - \exp(m_2)}, \qquad \begin{aligned} m_1 &= \epsilon^{-1}[-1 + \sqrt{(1 - \epsilon)}] = -\tfrac{1}{2} + O(\epsilon). \\ m_2 &= \epsilon^{-1}[-1 - \sqrt{(1 - \epsilon)}] = -2\epsilon^{-1} + O(1). \end{aligned}$$

The essential singularity at $\epsilon = 0$ is evident. Note that which boundary condition is lost as ϵ tends to zero, or, equivalently, which boundary is associated with the rapid variation of y making neglect of $\epsilon y''$ inappropriate, depends on the sign of ϵ.

$$\lim_{\epsilon \downarrow 0} y(t, \epsilon) = \exp[\tfrac{1}{2}(1 - t)], \qquad 0 < t \leqq 1;$$

$$\lim_{\epsilon \uparrow 0} y(t, \epsilon) \equiv 0, \qquad 0 \leqq t < 1.$$

This different behavior for different signs of ϵ is a reflection of the essential singularity in the solution. A striking illustration of this behavior in a boundary value problem for a partial differential equation can be found in [16]. An article by Erdelyi [4] provides a good introduction to recent work on singular perturbation problems.

Students sometimes remark that an asymptotic expansion valid as $\lambda \to \infty$ actually gives good answers for $|\lambda|$ as small as 5, say, while a convergent expansion about $\epsilon = 0$ only seems useful for numerical computation when $|\epsilon|$ is quite small, perhaps less than about 0.2. Since $1/5 = 0.2$ the remark as stated has little content. To give a concrete example, no superior merit can be assigned asymptotic approximations by noting that the leading term in the asymptotic expansion of $J_1(x)$ as $x \to \infty$,

$$J_1(x) \sim (2/\pi x)^{1/2} \cos\left(x - \frac{3\pi}{4}\right), \tag{6}$$

gives the first zero of 3.832 with an error of less than $2\frac{1}{2}\%$. When $x = (3.832)^{-1}$ the error in using the first term in the convergent expansion of $J_1(x)$ about $x = 0$ is less than 1%. The advantage of the asymptotic approximation (6) lies in its correct prediction of the decaying oscillatory behavior of J_1, a behavior which is not apparent from the series about the origin. (It is well known that the asymptotic power series for half-order Bessel functions terminate, thereby summing the convergent power series about $x = 0$. A recent example where an asymptotic attack on a nontrivial problem led to the same phenomenon, and hence an exact solution, can be found in [11].) The reward of obtaining a relatively simple expression giving the essential features of the solution comes to those who attack a problem from its most singular aspect. It appears that even in the impersonal

area of mathematical analysis there is an advantage in "taking the bull by the horns." This is simply because the behavior of a function is dominated by its worst singularity.

The above line of thought leads to an understanding of why asymptotic approximations are particularly important in applied mathematics. In seeking to understand a complex phenomenon the applied mathematician tries to select a simplified representation which retains its essentials. For example, in trying to understand phenomena arising from the hypersonic flight of a rocket, he may consider hypersonic flow past a sphere or a cone. Only in unusual instances would one be interested in quantitative aspects of the solution to such an idealized problem. As we have seen, if some sort of asymptotic analysis is possible, it would be expected to lead to a good qualitative picture of the solution. By using large computing machines one might well be able to get detailed information about hypersonic flow past a shape much more like that of a rocket than a sphere or a cone, but the problem now is more in the domain of the engineer than the applied mathematician. Moreover, one can hardly trust the result of a large detailed machine computation unless there is agreement with a basic understanding of the phenomenon, determined perhaps by a theoretician like an applied mathematician or perhaps by an experimental scientist.

It will be helpful if we illustrate our ideas by more specific remarks on another example. Consider the propagation of small amplitude water waves due to an initial displacement, at $t=0$, of height d for $|x| \leqq h$ and height zero for $|x| > h$. By use of simple Fourier integral techniques the elevation ζ can be expressed as

$$\zeta(x, t) = \frac{2d}{\pi} \int_0^\infty \sin kh \cos kx \cos \gamma t \frac{dk}{k}, \tag{7}$$

where γ is a function of k. For water of depth H, surface tension $T'\rho$, and density ρ,

$$\gamma^2 = (gk + T'k^3) \tanh kH, \tag{8}$$

where g is the acceleration of gravity [10, p. 512 and p. 515]. Equations (7) and (8) are a complete formal solution to the problem but one cannot be content to leave the answer in the form of a complicated definite integral. The reason for doing the problem is to learn something about water waves. About all it is possible to learn so far is how the phase velocity γ/k changes with surface tension and depth.

A simplification of (7) valid for small time can be obtained by expanding $\cos \gamma t$ in a Maclaurin series but we can guess in advance that this will not be very informative. The initially rectangular bump will simply spread out; by symmetry it must split into congruent right and left moving portions. The fact that the expansion for small t is convergent goes hand in hand with the fact that it expresses a relatively uninteresting regular perturbation from the initial state.

When t is large, we expect the effect of initial conditions to be small and the wave pattern to be dominated by whatever phenomena there are which are common to all water waves. The integrand of (7) has an essential singularity at $t=\infty$ so mathematical anticipation of an asymptotic expansion reinforces physical expectation that it will be most fruitful to examine (7) when t is large. The classical calculation requires investigating integrals for large t at fixed x/t by the method of steepest descent. It transpires that "periods, wave-lengths and wave-velocities are propagated with the group velocity [$d\gamma/dk$]; individual waves travel with the local wave-velocity [γ/k] but change their periods, lengths, and velocities as they travel," [10, p. 514]. Because of surface tension there is a minimum group velocity. Behind a point moving with this velocity the surface elevation falls off exponentially to the mean level. In the vicinity of this point, Airy functions are required mathematically. Airy functions are the simplest functions whose behavior changes from monotonic (describing smooth water behind the point moving with minimum group velocity) to oscillatory (describing waves in front of this point). There are larger longer gravity waves towards the front of the wave-train. Near a point moving with the maximum group velocity (which is the same as the phase velocity of pure gravity waves) there is again Airy function behavior as the surface elevation stops oscillating and exponentially approaches the mean level. Experimental verification that these predictions are correct and not very sensitive to initial conditions can be obtained by careful observation of what happens when one throws a rock into a pond.

It is clear that there is a wealth of valuable and elegant physical information concealed in (7) and revealed by an asymptotic approximation to (7). The concept of group-velocity is in fact central to an understanding of all types of wave-motion. This can be deduced from the fact that the nature of waves *in water* only appears in the dispersion relationship $\gamma=\gamma(k)$ of (8) and not in the Fourier integral (7), and the fact that qualitative behavior depends only on general properties like the existence of maxima and minima of $d\gamma/dk$. It is typical that a concept of great value in widely different contexts is easily illustrated in the context of fluid motion. The physical prerequisites are few and the phenomena are within the everyday experience of everyone. This is why study of subjects like fluid mechanics and elasticity usually form an important part of the basic education of an applied mathematician.

Returning to our example, in [1] there can be found an account of propagation of light signals. Included are subtle asymptotic mathematics, mastery of physical principles, and glimpses of human conflicts in the emergence of fundamental new ideas. Group velocity plays an important role, but a subtle one. W. Wien had noticed that if signals always travel with the group velocity then signals could travel faster than light. Because of a complex dispersion relation, Sommerfeld and Brillouin's resolution of this apparent conflict with relativity theory requires steepest descent calculations complicated by the presence of branch cuts and poles near the saddle points.

To repeat the qualification made at the beginning, in stressing the importance of asymptotics we do not mean to make an invidious comparison with other weapons of the applied mathematician not treated in this note. For example, only a few point masses need be involved in a gravitational interaction for useful results to be derivable by the asymptotic methods of statistical mechanics. Of great interest, however, are three-body problems where such many-particle methods are inapplicable. In any given situation, then, search for a "large" variable and an accompanying singular perturbation problem may be in vain. This is illustrated by the tale of the applied mathematician who resigned from a meteorology research group as soon as he discovered that all terms in the relevant equations were of the same order of magnitude. On the other hand, a skilled applied mathematician, aware of the power of asymptotic methods, can frequently formulate a problem in such a way that they can be used to reveal the essential features of the phenomenon (see [2]). One is reminded here of a remark concerning a certain distinguished scientist, "All he can do is solve boundary layer problems. . . . But of course he can turn all problems into boundary layer problems!"

To sum up, since he is most often interested in the qualitative behavior of the solution to an idealized problem, the applied mathematician frequently uses asymptotic expansions which give this behavior because they describe functions near their worst (essential) singularities. The applied mathematician has the technical ability to use asymptotic methods when the occasion demands. More important, he possesses the power of formulating problems so that asymptotic methods can be used when the occasion *permits*.

The author has been on leave, 1963–64, at the Massachusetts Institute of Technology, where he was partially supported by the National Science Foundation.

References

1. L. Brillouin, Wave Propagation and Group Velocity, Academic Press, New York, 1960.

2. G. Carrier, Boundary layer problems in applied mechanics, Advances in Applied Mechanics, (R. von Mises and T. von Karman, Eds.), Academic Press, New York, 3 (1953) 1–19.

3. A. Erdelyi, Asymptotic Expansions, Dover, New York, 1956.

4. ———, Singular perturbations of boundary value problems involving ordinary differential equations, J. Soc. Indust. Appl. Math., 11 (1963) 105–116.

5. A. Erdelyi and M. Wyman, The asymptotic evaluation of certain integrals, Arch. Rat. Mech. Anal., 14 (1963) 217–260.

6. K. O. Friedrichs, Asymptotic phenomena in mathematical physics, Bull. Amer. Math. Soc., 61 (1955) 485–504.

7. H. Greenberg, (Chmn.), Applied mathematics: what is needed in research and education, SIAM Review, 4 (1962) 297–319.

8. H. Greenspan, Applied mathematics as a science, this MONTHLY, 68 (1961) 872–880.

9. H. Jeffreys, Asymptotic Approximations, Clarendon Press, Oxford, 1962.

10. H. Jeffreys and B. Jeffreys, Methods of Mathematical Physics, 3rd ed., University Press, Cambridge, 1962.

11. H. A. Lang, Two verifications of the Karal-Keller theory of wave propagation, Proc. of the Fourth U. S. Congress of Applied Mechanics, Amer. Soc. Mech. Engrs., 1 (1962) 239–242.

12. P. Morse and H. Feshbach, Methods of Theoretical Physics, McGraw-Hill, New York, 1953.

13. F. J. Murray, Education for applied mathematics, this MONTHLY, 69 (1962) 347–357.

14. J. B. Rosser, Asymptotic formulas and series, Modern Mathematics for the Engineer (E. Beckenbach, Ed.), McGraw-Hill, New York, 1961, 133–163.

15. E. Titchmarsh, The Theory of Functions, 2nd Ed., University Press, Oxford, 1952.

16. W. Wasow, On the asymptotic solution of boundary value problems containing a parameter, J. Math. Phys., 23 (1944) 175–183.

A particularly illuminating and up-to-date account of singular perturbation theory can be found in M. Van Dyke's "Perturbation Methods in Fluid Mechanics," Academic Press, New York, 1964.

The referee has called my attention to a relevant article by M. Kruskal called "Asymptotology" ("the art of dealing with applied mathematical systems in limiting cases") in *Mathematical Models in Physical Sciences* (S. Drobot, Ed.), Prentice-Hall, Englewood Cliffs, N. J., 1963.

TWO CRITERIA FOR DEDEKIND DOMAINS

H. S. BUTTS AND L. I. WADE, Louisiana State University at Baton Rouge

1. Introduction. One of the purposes of classical algebraic number theory is the study of the ideal theory in the ring A of all algebraic integers of a finite algebraic extension of the field of rational numbers. Every proper ideal of A (that is, different from A and the zero ideal) is either a prime ideal or a (unique) product of prime ideals; furthermore the residue class ring (ring of cosets) of A modulo an ideal B is finite for every proper ideal B of A (see [4] p. 56). The number of elements in the residue class ring of A modulo the ideal B is called the norm of B and is denoted by $N(B)$. Using the unique factorization theorem stated above, it is shown that $N(BC)=N(B)\cdot N(C)$ for any two proper ideals B and C of A. This multiplicative property of the norm is a very important tool in classical algebraic number theory. The first main result of this paper is that the multiplicative property of the norm implies the unique (prime ideal) factorization theorem in an integral domain with unit element in which every proper ideal has a finite norm. The second main result is concerned with the notion of the length $L(B)$ of an ideal B, defined in the last section of the paper. It is shown in Theorem 4 that the unique (prime ideal) factorization theorem is valid in a domain with unit if and only if every proper ideal has finite length and $L(BC)=L(B)+L(C)$ for every pair of proper ideals B and C.

2. Preliminaries. We review some basic concepts.

Throughout this paper J will denote an integral domain with a multiplicative identity. An ideal B of J is a nonempty subset of J which is closed under subtraction and under multiplication by elements of J. If B and C are ideals in J, then the product BC is the set of all finite sums of the form $\sum_{i=1}^{n} b_ic_i$ where $b_i\in B$ and $c_i\in C$; the sum $B+C$ of B and C is the set of all elements of the form $b+c$ where $b\in B$ and $c\in C$. It is easy to show that sums and products of ideals are ideals and the usual commutative, associative, and distributive laws hold for

BC and $B+C$ and, moreover, that the zero ideal (0) and the domain J are the additive and multiplicative identities respectively. A proper ideal of a domain J is an ideal different from J and the zero ideal. An ideal N is called nonfactorable provided N is proper and $N=BC$, where B and C are ideals, implies that either $B=J$ or $C=J$. (In classical algebraic number theory this is the definition of "prime" ideal.) An ideal P is a prime ideal provided $ab\in P$ implies $a\in P$ or $b\in P$ where a and b are elements of J.

In a general integral domain a nonfactorable ideal may not be a prime ideal and a prime may not be nonfactorable. (See examples 5 and 6 in section 4.) An ideal M in a domain J is said to be maximal provided $M\neq J$ and the only ideals of J which contain M are M and J. It is a well-known fact that in a domain with unit element, maximal ideals are prime (but not conversely). An ideal B is said to be finitely generated if there exist elements $b_i\in B$ for $i=1, \cdots,$ n such that every element of B is expressible in the form $\sum_{i=1}^{n} x_i\cdot b_i$ with $x_1, \cdots, x_n$ elements of J; in this case we write $B=(b_1, \cdots, b_n)$ and say that B is generated by $b_1, \cdots, b_n$.

A domain J with unit is called a Dedekind domain provided every proper ideal is either a prime ideal or a unique (except for order) product of prime ideals. It can be shown (see [1] p. 273) that a domain with unit, in which every proper ideal is either a prime ideal or a product of prime ideals, is a Dedekind domain and that proper prime ideals are maximal in a Dedekind domain. The domains in the following three examples are Dedekind domains having the finite norm property (i.e., the ring of cosets modulo any proper ideal is finite).

Example 1. The ring of all algebraic integers in any finite algebraic extension of the field of rational numbers (in particular, the ring of "ordinary" integers) is a Dedekind domain having the finite norm property (see [4]).

Example 2. The ring of polynomials $F[x]$ over a finite field F is a Dedekind domain which has the finite norm property (see [5] pp. 55–62).

Example 3. Let p be a prime number and denote by D the set of rational numbers a/b, where a and b are integers with b not divisible by p. The ideal P generated by p in D is the only proper prime ideal in D and every proper ideal in D is a power of P. The domain D is a Dedekind domain having the finite norm property (see [1] pp. 223–230).

3. Finite norm. In this section we assume that the domain J has the property that every proper ideal B has finite norm $N(B)$, that is, the residue class ring of J modulo B is finite with $N(B)$ elements.

LEMMA 1. *If B and C are proper ideals of J such that $B\subset C$, then there exists a positive integer k such that $N(B)=kN(C)$. Furthermore, B is a proper subset of C if and only if $k\geqq 2$.*

Proof. Since B is a subgroup of C under addition it follows easily that there exist elements $c_i\in C$ for $i=1, \cdots, k$ such that the cosets $\{B+c_i\}$ are disjoint for $i=1, \cdots, k$ and $C=\bigcup_{i=1}^{k}\{B+c_i\}$. It is clear that if $B=C$ then $k=1$, and

if $B \neq C$ then $k > 1$. If $J = \bigcup_{j=1}^{n} \{C + x_j\}$ is the coset decomposition of J modulo C, then

$$J = \bigcup_{\substack{1 \leq j \leq n \\ 1 \leq i \leq k}} \{B + c_i + x_j\}$$

is the coset decomposition of J modulo B. It follows that $N(B) = kN(C)$.

LEMMA 2. *If every proper ideal of J has finite norm, then every proper ideal is finitely generated.*

Proof. Let B be a proper ideal of J and $b_1 \neq 0$ an element of B. If $B = (b_1)$, then B is finitely generated. If $B \neq (b_1)$, let $b_2 \in B$ and $b_2 \notin (b_1)$. If $B = (b_1, b_2)$ then B is finitely generated. If $B \neq (b_1, b_2)$, then there exists $b_3 \in B$ such that $b_3 \notin (b_1, b_2)$. If this process continues, we obtain a chain of ideals

$$(b_1) \subset (b_1, b_2) \subset (b_1, b_2, b_3) \subset \cdots.$$

By Lemma 1 we have

$$N((b_1)) > N((b_1, b_2)) > N((b_1, b_2, b_3)) > \cdots$$

and since the norm of a proper ideal is a positive integer, it follows that the chain of ideals above must terminate in a finite number of steps and, consequently, B is finitely generated.

THEOREM 1. *Let J have the property that $N(AB) = N(A)N(B)$ for every pair of proper ideals A and B in J. If A and B are ideals in J such that $A \subset B$, then there exists an ideal C in J such that $A = BC$.*

Proof. We first prove the theorem for the case in which $B = A + (b)$ where $b \notin A$. If either A or $A + (b)$ is not a proper ideal or if $b \in A$, then it is obvious that there is an ideal C such that $A = BC$. Suppose that A and $B = A + (b)$ are proper ideals such that $b \notin A$ and denote by C the set of elements $x \in J$ for which $xb \in A$. It is easy to verify that C is an ideal and that $BC \subseteq A$. In order to show that $BC = A$, it is sufficient (by Lemma 1) to prove that $N(BC) = N(A)$. Since $BC \subseteq A$ it follows from Lemma 1 that $N(BC) = N(B) \cdot N(C) \geq N(A)$. Therefore, in order to prove that $N(BC) = N(A)$, it is sufficient to show that $N(BC) = N(B)N(C) \leq N(A)$. Since $A \subset B$ it follows (by Lemma 1) that $N(A) \div N(B) = k$ is an integer. Hence it is sufficient to show that $N(C) \leq k$.

Let $B = \bigcup_{i=1}^{k} \{A + x_i\}$ be a coset decomposition of B modulo A and let $J = \bigcup_{i=1}^{n} \{C + y_i\}$ be a coset decomposition of J modulo C, where $k = N(A) \div N(B)$ and $n = N(C)$. We will show that the set of cosets $\{C + y_1, C + y_2, \cdots, C + y_n\}$ can be mapped one-to-one onto a subset of the set of cosets $\{A + x_1, A + x_2, \cdots, A + x_k\}$, from which it will follow that $N(C) = n \leq k$. Associate with each coset $C + y_i$ the coset $A + x_j$ in which the element by_i occurs (where $B = A + (b)$). If $C + y_r = C + y_s$, then $y_r - y_s \in C$. From the definition of the ideal C it follows that $(y_r - y_s)b \in A$. Since $y_r b$ and $y_s b$ belong to the ideal B and

$y_r b - y_s b \in A$ it follows that $y_r b$ and $y_s b$ belong to the same coset of B modulo A. This shows that the mapping defined above is independent of the coset representative. Since each element of B belongs to one and only one coset of A, the mapping is single valued. Suppose that two cosets, $C+y_u$ and $C+y_v$, are associated with the same coset $A+x_j$, so that $by_u - by_v \in A$. It follows that $b(y_u - y_v) \in A$. Therefore $y_u - y_v \in C$ and $C+y_u = C+y_v$. Hence $N(C) = n \leqq k$ and $A = BC$ in case $B = A+(b)$.

Now let A and B be any two proper ideals such that $B \supset A$. By Lemma 2, B is finitely generated and, consequently, there exist elements $b_1, b_2, \cdots, b_n$ such that

$$A \subset A + (b_1) \subset A + (b_1) + (b_2) \subset \cdots \subset A + (b_1) + \cdots + (b_n) = B.$$

Applying the result obtained above in the first part of this proof, we obtain an ideal C_1 such that $A = [A+(b_1)]\cdot C_1$. There is an ideal C_2 such that $A+(b_1) = [A+(b_1)+(b_2)]\cdot C_2$. Therefore $A = [A+(b_1)+(b_2)]\cdot C_1\cdot C_2$. By a standard induction argument it follows that there exists an ideal C such that $A = BC$.

LEMMA 3. *If P is a prime ideal in any domain and A and B are ideals such that $AB \subset P$, then either $A \subset P$ or $B \subset P$.*

Proof. Suppose $A \not\subset P$ and let $a \in A$, $a \notin P$. Since $AB \subset P$, $ab \in P$ for every $b \in B$. It follows from the definition of a prime ideal that $b \in P$ for every $b \in B$.

LEMMA 4. *If B and C are ideals in a domain D with unit element and $0 \neq a \in D$, then $(a)B = (a)C$ implies $B = C$.*

Proof. If $b \in B$ then $ab \in (a)C$ and therefore

$$ab = \sum_{i=1}^{n} (x_i a)c_i = \sum_{i=1}^{n} a(x_i c_i) = \sum_{i=1}^{n} ac_i' = a\left(\sum_{i=1}^{n} c_i'\right),$$

where $x_i \in D$, $c_i \in C$, $c_i' \in C$ for $i = 1, \cdots, n$. Therefore

$$ab = a\left(\sum_{i=1}^{n} c_i'\right) \quad \text{and} \quad b = \sum_{i=1}^{n} c_i' \in C.$$

THEOREM 2. *If J has the property that $N(AB) = N(A)N(B)$ for every pair of proper ideals A and B in J, then every proper ideal of J is either a prime ideal or a unique (except for order) product of prime ideals.*

Proof. Let A be any proper ideal of J. If A is nonfactorable then A is maximal (and hence prime) since $A \subset B$ implies by Theorem 1 that there is an ideal C such that $A = BC$, from which it follows that either B or C is J and therefore B is either A or J. If A is not nonfactorable, then $A = A_1 A_2$ where A_1 and A_2 are proper ideals. Hence $N(A) = N(A_1)N(A_2)$. Since the norm of a proper ideal is an integer $\geqq 2$ it follows that A can be factored into a product of nonfactorable ideals, say $A = N_1 \cdot N_2 \cdot \cdots \cdot N_n$, where N_i is a proper nonfactorable ideal for

$i=1, \cdots, n$. Since proper nonfactorable ideals are maximal, it follows that A is a product of maximal (hence prime) ideals.

Uniqueness follows from [1] page 273, but a much more elementary proof can be given in this case. Suppose that

$$A = M_1 M_2 \cdots M_r = M_1' M_2' \cdots M_s',$$

where the M_i are maximal and the M_i' are proper prime ideals. By Lemma 3, $M_j \subset M_1'$ for some j, say $j=1$. Then $M_1 = M_1'$. Let $x \neq 0$ belong to M_1. Since $(x) \subset M_1$, there exists an ideal Q such that $(x) = M_1 Q$ by Theorem 1. We have

$$QM_1M_2 \cdots M_r = QM_1M_2' \cdots M_s'$$

$$(x) \cdot M_2 \cdot \cdots \cdot M_r = (x) \cdot M_2' \cdot \cdots \cdot M_s'$$

By Lemma 4, we have $M_2 \cdot M_3 \cdot \cdots \cdot M_r = M_2' \cdot M_3' \cdot \cdots \cdot M_s'$.

A standard induction argument completes the proof.

We remark that in a domain J with unit element in which proper ideals have finite norms, the converse to Theorem 2 is also valid. The standard proof in classical algebraic number theory is all that is needed to get this result. A domain with unit element in which the unique prime ideal factorization theorem holds need not, however, have the finite norm property, as the following example shows.

Example 4. The ring of polynomials $F[x]$ over an infinite field F is a Dedekind domain which does not have the finite norm property, since the residue class ring of $F[x]$ modulo the ideal (x) is isomorphic to F.

4. Chains of ideals. In this section we consider a general integral domain J with unit element. Let A be a proper ideal of J and let S be the set of all positive integers n obtained by taking the lengths of all possible finite chains of ideals

$$A \subset A_1 \subset A_2 \subset \cdots \subset A_n = J, \quad \text{where} \quad A \neq A_1 \text{ and } A_i \neq A_{i+1}$$

for $i=1, 2, \cdots, n-1$. We say that A has finite length in case the set S is finite and, in this case, the largest integer in S will be called the length of A and denoted by $L(A)$.

LEMMA 5. *If every proper ideal in J has finite length, then every proper ideal is finitely generated.*

Proof. Let $a_1 \neq 0$ be an element of the proper ideal A. If $A = (a_1)$ then A is finitely generated. If $A \neq (a_1)$, let $a_2 \in A$, $a_2 \notin (a_1)$. If $A = (a_1, a_2)$, then A is finitely generated. If $A \neq (a_1, a_2)$, let $a_3 \in A$ and $a_3 \notin (a_1, a_2)$. The chain of ideals $(a_1) \subset (a_1, a_2) \subset (a_1, a_2, a_3) \subset \cdots$ must terminate since (a_1) has finite length, and it follows that A is finitely generated.

THEOREM 3. *Let the integral domain J have the property that every proper ideal has finite length and that $L(AB) = L(A) + L(B)$ for every pair of proper ideals A and B in J. If A and B are ideals in J such that $A \subset B$, then there exists an ideal C such that $A = BC$.*

Proof. As in the proof of Theorem 1, it is sufficient to prove the theorem for the case in which $B=A+(b)$, where $b\notin A$. Denote by C the set of elements x in J for which $xb\in A$. As in Theorem 1, C is an ideal and $BC\subset A$. Hence $L(BC)=L(B)+L(C)\geqq L(A)$. In order to prove that $BC=A$ it is sufficient to prove that $L(A)\geqq L(BC)=L(B)+L(C)$.

Let

$$(1)\quad C\subset C+(x_1)\subset C+(x_1)+(x_2)\subset\cdots\subset C+(x_1)+\cdots+(x_n)=J$$

be a strictly increasing sequence of ideals from C to J. We will show that the sequence

$$(2)\quad A\subset A+(bx_1)\subset A+(bx_1)+(bx_2)\subset\cdots\subset A+(bx_1)+\cdots+(bx_n)$$

is a strictly increasing sequence of ideals each contained in B. It will follow from this that $L(A)\geqq L(B)+L(C)$ and therefore $A=BC$.

Since $B=A+(b)$, it is clear that each of the ideals in the sequence (2) is contained in B. If $A=A+(bx_1)$, then $bx_1\in A$ and it follows from the definition of C that $x_1\in C$. Hence $C=C+(x_1)$ and the sequence (1) is not strictly increasing. A standard induction argument completes the proof.

It is known that a domain J with unit is Dedekind if and only if $B\supset A$ implies there exists an ideal C such that $A=BC$ (see [2] p. 13). We will give a more elementary proof that the domain of Theorem 3 is Dedekind.

LEMMA 6. *If $a\in J$ and B is an ideal in J such that $(a)\supset B$, then there exists an ideal Q in J such that $(a)\cdot Q=B$.*

Proof. Denote by Q the set of elements $x\in J$ such that $ax\in B$. It is easy to show that Q is an ideal and $(a)Q=B$.

LEMMA 7. *Let J have the property that every proper ideal in J is either maximal or a product of maximal ideals. If M is a maximal ideal in J, then there exists an element $a\in J$ and an ideal $Q\neq(0)$ in J such that $MQ=(a)$.*

Proof. If $M=(0)$, take $Q=J$ and $a=0$. If $M\neq(0)$, let $a\neq 0$ be an element of M. There exist maximal ideals $M_1, M_2,\cdots, M_n$ in J such that $(a)=M_1M_2\cdots M_n$. Since $M\supset(a)$, it follows from Lemma 3 that $M=M_i$ for some i, say $i=1$. Then $(a)=MM_2\cdots M_n$ and set $Q=M_2\cdot\cdots\cdot M_n$.

LEMMA 8. *Let J have the property that every proper ideal in J is either maximal or a unique (except for order) product of maximal ideals. If A and B are proper ideals in J such that $A\subset B$, then there exists an ideal C in J such that $A=BC$.*

Proof. By Lemma 7 there exists an ideal $Q\neq(0)$ in J and an element $b\in J$ such that $BQ=(b)$. It follows that $(b)=BQ\supset AQ$ and by Lemma 6 there exists an ideal C in J such that $(b)C=AQ$. Then $(b)C=BQC=AQ$. By factoring A, B, C and Q into maximal ideals and using the uniqueness of the factorization, we see that $A=BC$.

THEOREM 4. *In order that J have the property that every proper ideal is either maximal or a unique (except for order) product of maximal ideals it is necessary and sufficient that every proper ideal have finite length and that $L(AB)=L(A)+L(B)$ for every pair of proper ideals A and B in J.*

Proof. In order to prove the necessity it is sufficient to note that if A is a proper ideal of J and $A=M_1M_2 \cdots M_n$ is the maximal ideal decomposition of A, then $L(A)=n$; this follows easily from Lemma 8 and the uniqueness of the factorization into maximal ideals.

The sufficiency follows from the fact that every proper ideal of J is finitely generated and there are no ideals properly between a maximal ideal and its square (see [3] pp. 33), however, a much more elementary proof will be given here.

Let A be a proper ideal of J. If A is nonfactorable, then it follows from Theorem 3 that A is maximal. If A is not nonfactorable, then proper ideals A_1 and A_2 exist such that $A=A_1A_2$. Since $L(A)=L(A_1)+L(A_2)$ it follows easily that A is a product of nonfactorable ideals. By Theorem 3, nonfactorable ideals are maximal and therefore A is a product of maximal ideals. The uniqueness follows as in the proof of Theorem 2.

The following two examples illustrate some of the relationships between various properties of ideals in a domain.

Example 5. Let $F[x]$ be the ring of polynomials over a finite field F and denote by D the subring of $F[x]$ consisting of all polynomials with x-coefficient zero. The domain D has the finite norm property, but the norm is not multiplicative and D is not a Dedekind domain. Every ideal in D is generated by either one or two elements and proper prime ideals are maximal. Every proper ideal is a product of nonfactorable ideals, but not uniquely (see [6]). The ideal (x^2) is a nonfactorable ideal and not a prime ideal (or, a product of prime ideals). The cancellation law for ideals is not valid in D. In D every ideal has finite length, but it is not true that $L(BC)=L(B)+L(C)$ for every pair of proper ideals B and C in D.

Example 6. Let $R[x, y]$ be the ring of polynomials in two variables over the field of rational numbers R and define a function v from $R[x, y]$ into the additive group of real numbers by

$$v\left(\sum_{i,j} a_{ij}x^iy^j\right) = \min_{a_{ij}\neq 0} \{i+j\sqrt{2}\}$$

for nonzero polynomials and define $v(0)=\infty$ (in the extended reals). Extend this function v from $R[x, y]$ to the quotient field $R(x, y)$ of $R[x, y]$ by $v(f/g)=v(f)-v(g)$. Denote by D the set of all $r\in R(x, y)$ such that $v(r)\geqq 0$ and by M the set of all $r\in R(x, y)$ such that $v(r)>0$. It can be shown that D is a domain with unit and that M is a maximal (proper, prime) ideal of D and $M^2=M$, i.e., M is a prime ideal which is not nonfactorable. The function v is a nondiscrete (rank one) valuation of $R(x, y)$ and D is the valuation ring of v (see [5] pp. 239–240).

The ideal M is not finitely generated and D does not have the finite norm property or the finite length property (see Lemma 2 and Lemma 5).

References

1. O. Zariski and P. Samuel, Commutative Algebra, vol. 1, Van Nostrand, Princeton, N.J. 1958.

2. W. Krull, Idealtheorie, Chelsea, New York, 1948.

3. I. S. Cohen, Commutative rings with restricted minimum condition, Duke Math. J., 17 (1950).

4. H. B. Mann, Introduction to Algebraic Number Theory, The Ohio State University Press, 1955.

5. B. L. van der Waerden, Modern Algebra, Ungar, New York, 1949.

6. H. S. Butts, Unique factorization of ideals into nonfactorable ideals, Proc. Amer. Math. Soc., 15 (1964).

EXTENDED NIELSEN OPERATIONS IN FREE GROUPS

J. J. ANDREWS, Florida State University, and M. L. CURTIS, Rice University

1. Introduction. The authors have made a conjecture about free groups [1] and noted some interesting topological consequences which would follow from the truth of the conjecture. The present paper represents some partial results obtained in attempting to prove the conjecture. Proposition 3.2 is due to Peter Hilton.

Let $X=\{x_1, \cdots, x_n\}$ be a finite set. Let W be the set of all monomials

$$x_{i_1}^{n_i} x_{i_2}^{n_2} \cdots x_{i_e}^{n_e},$$

where $1 \leqq i_j \leqq n$, n_i is an integer, and $e \geqq 0$. (For $e=0$ we write 1.) The product of two elements

$$w_1 = x_{i_1}^{n_1} \cdots x_{i_e}^{n_e}, \qquad w_2 = x_{j_1}^{m_1} \cdots x_{j_k}^{m_k}$$

is defined by the following:

$$w_1 \cdot w_2 = x_{i_1}^{n_1} \cdots x_{i_e}^{n_e} x_{j_1}^{m_1} \cdots x_{j_k}^{m_k}.$$

Under this operation W forms a semi-group with unit. We identify two elements w_1, w_2 of W if one can be obtained from the other by

(1) replacing $x_i^m x_i^n$ by x_i^{m+n}; and (2) replacing x_i^0 by 1

or the inverse of (1) and (2).

The resulting set forms a group F. The group F is said to be free on $x_1, \cdots,$ x_n. We say F is free on $y_1, \cdots, y_n$ if there is an isomorphism $\theta: F \rightarrow F$ such that $\theta(x_i) = y_i$.

Let F be a free group on $x_1, \cdots, x_n$. If $a_1, \cdots, a_p$ is any finite set of elements of F, then a *Nielsen operation* is defined to be a change from $a_1, \cdots, a_p$ to $\bar{a}_1, \cdots, \bar{a}_p$ in one of the following three ways:

(i) $\bar{a}_1 = a_1^{-1}, \bar{a}_2 = a_2, \cdots, \bar{a}_p = a_p$,
(ii) $\bar{a}_1 = \bar{a}_i, a_i = a_1, \bar{a}_j = a_j$ for $j \neq 1, i$.
(iii) $\bar{a}_1 = a_1 a_2, \bar{a}_2 = a_2, \cdots, \bar{a}_p = a_p$.

THEOREM (NIELSEN [2]). *If F is free on $x_1, \cdots, x_n$ and also free on $y_1, \cdots, y_n$, then a finite sequence of Nielsen operations will change $x_1, \cdots, x_n$ to $y_1, \cdots, y_n$. (That is, Nielsen operations generate the group of automorphisms of F.)*

If A is a subset of F then $G(A)$ will denote the subgroup of F generated by A and $N(A)$ will denote the normal subgroup of F generated by A. If A is a finite set, then we note that Nielsen operations on A do not change $G(A)$. Given a presentation

$$\{x_1, \cdots, x_n \mid r_1, \cdots, r_n\}$$

of the trivial group, we know only that $N(r_1, \cdots, r_n) = F$ and it often happens that $G(r_1, \cdots, r_n) \neq F$. In these cases Nielsen operations cannot change $r_1, \cdots, r_n$ to $x_1, \cdots, x_n$, but we conjecture that addition of a fourth type of operation suffices.

(iv) $\bar{a}_1 = g a_1 g^{-1}, \bar{a}_2 = a_2, \cdots, \bar{a}_p = a_p$, where g is any element of F.

The three Nielsen operations along with (iv) will be referred to as *extended Nielsen operations*. The symbols

$$(r_1, \cdots, r_n) \xrightarrow{N} (x_1, \cdots, x_n)$$

will mean that $r_1, \cdots, r_n$ may be changed to $x_1, \cdots, x_n$ by a finite sequence of extended Nielsen operations.

CONJECTURE. *If F is free on $x_1, \cdots, x_n$ and $N(r_1, \cdots, r_n) = F$, then*

$$(r_1, \cdots, r_n) \xrightarrow{N} (x_1, \cdots, x_n).$$

Just as the three Nielsen operations on A do not change $G(A)$, so the extended Nielsen operations on A do not change $N(A)$. Hence, the condition $N(r_1, \cdots, r_n) = F$ is surely necessary for $(r_1, \cdots, r_n) \xrightarrow{N} (x_1, \cdots, x_n)$, so the conjecture could have been stated with "if and only if." We will see that $G(A)$ may be changed by extended Nielsen operations—otherwise the conjecture would be clearly false.

We restrict attention to the case $n = 2$, primarily for notational simplicity. It is not clear, however, that a proof for $n = 2$ would give a means of proving the conjecture for larger values of n.

2. Degrees of words. In the free group $F(a, b)$ on a and b, we define the *a-degree* $d_a(w)$ of a word $w \in F$ to be the sum of all the exponents of a in w. For any pair (p, q) of integers there is a normal subgroup consisting of all words w such that $d_a(w)$ is a multiple of p and $d_b(w)$ is a multiple of q. Clearly, such a subgroup is normal. For $(p, q) = (0, 0)$ it is the commutator subgroup.

PROPOSITION 2.1. *$c \in F$ belongs to $F' = [F, F]$ if and only if $d_a(c) = 0 = d_b(c)$.*

Proof. For $c \in F'$ it is obvious that both degrees are zero. Conversely, if $d_a(c) = 0 = d_b(c)$ then c goes to $0a + 0b$ in the abelianization of F. Hence $c \in F'$.

We are interested in pairs (r, s) of elements of F whose normal closure is F. This imposes a condition on the degrees.

PROPOSITION 2.2. *If $N(r, s) = F$, then*

$$\begin{vmatrix} p & q \\ l & m \end{vmatrix} = \pm 1,$$

where

$$d_a(r) = p, \qquad d_b(r) = q,$$
$$d_a(s) = l, \qquad d_b(s) = m.$$

Proof. It suffices to look at the abelianized group. Then a, b are linear integral combinations of r, s (actually we mean their images in the free abelian group). So there exists a matrix

$$\begin{pmatrix} \rho & \sigma \\ t & u \end{pmatrix}$$

of integers such that

$$\begin{pmatrix} \rho & \sigma \\ t & u \end{pmatrix}\begin{pmatrix} p & q \\ l & m \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$\begin{vmatrix} \rho & \sigma \\ t & u \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} p & q \\ l & m \end{vmatrix}$$

are integers whose product is 1, and the proposition is proved.

This proposition is not the best possible. If we require only that $\{a, b \mid r, s\}$ be a perfect group (see Section 3), then the same result may be proved.

PROPOSITION 2.2a. *If $F/N(r, s)$ is equal to its commutator subgroup, then*

$$\begin{vmatrix} p & q \\ l & m \end{vmatrix} = \pm 1.$$

Of course, the converse of Proposition 2.2 is far from true (Section 3 contains specific examples), but there is one special case in which it is.

THEOREM 2.1. *If* $r=a^pb^q$, $s=a^lb^m$, *then*

$$N(r, s) = F \Leftrightarrow \begin{vmatrix} p & q \\ l & m \end{vmatrix} = \pm 1 \Leftrightarrow (r, s) \overset{N}{\to} (a, b).$$

Proof. We have observed in Section 1 that none of the extended Nielsen operations changed the normal closure of a set, so always

$$(r, s) \overset{N}{\to} (a, b) \Rightarrow N(r, s) = F.$$

Hence, it suffices to prove that the condition on the determinant implies $(r, s)\overset{N}{\to}(a, b)$. This is an elementary manipulative exercise, a procedure for which we now sketch.

Suppose $p \geqq l$. Then r is changed to $a^lb^qa^{p-l}$ (by conjugation by a^{-p+l}) and then changed to the inverse $a^{-p+l}b^{-q}a^{-l}$ of this element. Then we multiply by s on the right. This gives

$$a^{-p+l}b^{-q+m} \quad \text{and} \quad a^lb^m.$$

Next we reduce b^m in the way we did a^p, etc. This will finally give a and b.

3. Perfect groups. If $N(r, s)=F$ then under abelianization $\eta: F \to F/F'$ we see that $\eta(r)$, $\eta(s)$ are a basis for the free abelian group F/F'. Hence $\eta(r)$, $\eta(s)$ may be changed by Nielsen operations only to $\eta(a)$, $\eta(b)$. By simply following these operations up in F we get the first result of this section.

PROPOSITION 3.1. *If* $N(r, s)=F$, *then* (r, s) *may be changed by Nielsen operations (no conjugation needed) to* $(\bar{r}, \bar{s})$ *such that* $\bar{r}=c_1a$, $\bar{s}=c_2b$, *with* $c_1, c_2 \in F'$.

It follows that we can restrict consideration to pairs of this sort, but it turns out that we cannot hope to handle all such pairs because the converse of Proposition 3.1 is false. That is, there exist elements c_1, c_2 of the commutator subgroup F' such that

$$N(c_1a, c_2b) \neq F.$$

Example. In particular, we assert that

$$c_1 = (a^{-1}bab^{-1})[(ba^{-1}b^{-1}a)(ab^{-1}a^{-1}b)]^3$$
$$c_2 = (ba^{-1}b^{-1}a)(ab^{-1}a^{-1}b)$$

is such a pair. In other words, the group G with presentation

$$G = \{a, b \mid c_1a, c_2b\}$$

is nontrivial.

To establish the nontriviality we will give a homomorphic mapping ϕ of G

onto A_5, the alternating group on five objects. Let

$$\phi(a) = (2\ 1\ 4)$$
$$\phi(b) = (1\ 2\ 3\ 4\ 5),$$

and one can check that $\phi(c_1a) = 1 = \phi(c_2b)$ so that this assignment does determine a homomorphism of G. Clearly, the image is nontrivial. (Actually the homomorphism is onto since (2 1 4) and (1 2 3 4 5) generate A_5.)

PROBLEM 1. In $F(a, b)$ describe those pairs $c_1, c_2 \in F'$ such that $N(c_1a, c_2b) = F$.

A good clue to this problem is furnished by an observation of Peter Hilton showing that the bad pairs $c_1, c_2(N(c_1a, c_2b) = F)$ come from presentations of perfect groups.

Recall that a group G is perfect if $G = G'$, its commutator subgroup. We note that for any $c_1, c_2 \in F'$ the presentation

$$\{a, b \mid c_1a, c_2b\}$$

is always a perfect group, since a group is perfect if and only if its abelianization is trivial.

PROPOSITION 3.2. *Each presentation of a nontrivial perfect group leads to a bad pair, $c_1, c_2 \in F'$. (This is not precise. We should say each presentation with two generators, but as remarked in Section 1, we use two generators only for notational convenience.)*

Proof. Let $G = \{a, b \mid r, s\}$ be perfect. Then a, b may be written as products of commutators of words in a and b (Proposition 2.1 applies only to free groups),

$$a = c_1(a, b)$$
$$b = c_2(a, b).$$

The natural map η: $F(a, b) \to G$ sends the word $c_1(a, b)$ to a in G. Hence $N(c_1^{-1}a, c_2^{-1}b)$ is in the kernel of η. It follows that $N(c_1^{-1}a, c_2^{-1}b) \neq F$.

So we see that one answer to Problem 1 is that c_1, c_2 is a bad pair if and only if c_1^{-1}, c_2^{-1} give a and b in some perfect group generated by a and b. This is not much help since no scheme for listing all finitely-generated perfect groups is available.

4. Lengths of words. In this section we consider the most promising tool for proof of our conjecture. The *length* $|w|$ of a word w in $F(a, b)$ is simply the number of letters used in writing it in reduced form (i.e., all aa^{-1} etc. are cancelled out). Obviously, for any two words $u, v \in F$

$$|u\ v| \leqq |u| + |v|$$

with equality holding whenever the last letter of a is not the inverse of the first letter of v.

In general, long words do not generate as much of F as short words, and the

same is true about their normal closure. When we are trying to change r, s to a, b by extended Nielsen operations, it seems natural to try to make $|r|+|s|$ decrease. Let us define $|r|+|s|$ to be *minimal* if

$$r, s \xrightarrow{N} (\bar{r}, \bar{s}) \Rightarrow |\bar{r}|+|\bar{s}| \geqq |r|+|s|.$$

We conjecture that if $N(r, s)=F$ and $|r|+|s|$ is minimal, then $|r|+|s|=2$. This, of course, would imply our original conjecture of section 1. In this section we give some very partial results along this line.

LEMMA 4.1. *$|r|+|s|$ minimal$\Rightarrow|r^2|=2|r|$, $|s^2|=2|s|$.*

Proof. If $|rr|<|r|+|r|$ then some cancellation has occurred, so that $r=pqp^{-1}$ with $|p|\geqq 1$. But then $(p^{-1}rp, s)$ is shorter than (r, s) (precisely, $|p^{-1}rp|+|s|<|r|+|s|$), contradicting the minimality of (r, s).

LEMMA 4.2. *$|r|+|s|$ minimal$\Rightarrow r$, s cyclically reduced.*

Proof. Cyclically reduced means (reduced and) the first and last letters are not conjugates. The proof is obvious.

LEMMA 4.3. *$|r|+|s|$ minimal$\Rightarrow r$ and s considered as cyclic words contain no common word w (or w and w^{-1}) with $|w|>\frac{1}{2}$ min $\{|r|, |s|\}$.*

Proof. Suppose $r=pwq$, $s=twu$. Then

$$\begin{array}{ccl} & & \overset{N}{\nearrow}\ (t^{-1}u^{-1}qp,\ t^{-1}u^{-1}w^{-1}) \\ (r, s) \xrightarrow{N} (wqp,\ t^{-1}u^{-1}w^{-1}) & N & \qquad \text{or} \\ & & \searrow\ (wqp,\ t^{-1}u^{-1}qp). \end{array}$$

Suppose $|r|\leqq|s|$. Then consider the lower pair.

$$\begin{aligned} |wqp| + |t^{-1}u^{-1}qp| &= |r| + |s| - |w| + |r| - |w| \\ &< |r| + |s| - \tfrac{1}{2}|r| + |r| - \tfrac{1}{2}|r| = |r| + |s|, \end{aligned}$$

contradicting the minimality of $|r|+|s|$. The same argument clearly applies if one (of r and s) contains w and the other w^{-1}.

The hypothesis that $N(r, s)=F$ means that a (and b) can be written as a product of conjugates of powers of r and s

$$a = f_1\beta^{l_1}f_1^{-1}\cdots f_k\beta^{l_k}f_k^{-1},$$

where $\beta=r$ or s, $l_1, \cdots, l_k$ are integers and $f_1, \cdots, f_k$ are some elements of F. In other words, this complicated expression cancels down to a by simply cancelling a and a^{-1}, b and b^{-1} whenever they become adjacent. We want to show that this expression collapses as we make $|r|+|s|$ minimal.

A simple result in this direction is the following.

LEMMA 4.4. *$|r|+|s|$ minimal and $a=frf^{-1}ese^{-1}$ imply that $|r|+|s|=2$. (Hence, if $N(r, s)=F$, r and s are the set $\{a, b\}$).*

Proof:

$$\begin{array}{ccl} & & N\ (a, s) \\ & \nearrow & \\ (r, s)\ N & & \\ & \searrow & \\ & & (r, a) \end{array}$$

so that

$$\begin{aligned} |r| + |s| &\leqq 1 + |s| \\ |r| + |s| &\leqq |r| + 1. \end{aligned}$$

By elementary considerations of degrees we see that one cannot have

$$a = frf^{-1}ere^{-1} \quad \text{or} \quad a = trt^{-1}ur^{-1}u^{-1}.$$

Let us agree to call a minimal pair (r, s) *nontrivial* if $|r|+|s|>2$. We have now shown the following.

THEOREM 4.1. *Neither a nor b can be expressed as a product of two conjugates of elements from a non-trivial minimal set.*

REMARK. With a little more effort the word "two" in Theorem 4.1 can be changed to "three," but it is not clear how to make a general inductive step.

If (r, s) may be changed so that one of them becomes quite short (and $N(r, s)=F$), then they may be changed to (a, b).

THEOREM 4.2. *If $N(r, s)=F$ and $|r|<4$, then $(r, s)\xrightarrow{N}(a, b)$.*

Proof. Since r is cyclically reduced and r cannot be a power of a or b except the first power (by elementary degree considerations), we see that the only possibilities are

$$a,\ b,\ ab,\ ab^{-1},\ a^2b,\ a^{-2}b,\ ab^2,\ ab^{-2},$$

or inverses of these. Again elementary degree considerations show that r and s always have a common subword of length $\geqq\frac{1}{2}|r|$. Hence $|r|+|s|$ may be reduced to 2, and the theorem follows.

It is clear that the arguments of Theorem 4.2 go just as well if a and b are replaced by any two words α, β such that $G(\alpha, \beta)=F$. This observation leads to another special case in which our conjecture is true.

THEOREM 4.3. *If $c\in F'$, then $(c^pa, c^qb)\xrightarrow{N}(a, b)$ for any integers p, q.*

Proof. We will show that if α, β are words such that $G(\alpha, \beta)=F$ and if $|p|\geqq|q|>0$, then

$$(c^p\alpha, c^q\beta) \xrightarrow{N} (c^{p'}\alpha', c^q\beta'),$$

where $|p'| < |p|$ and $G(\alpha', \beta') = F$. This will suffice to prove the theorem, since $(c^r\alpha, \beta) \xrightarrow{N} (\alpha, \beta)$.

But we see easily that, for $q > 0$,

$$(c^p\alpha,\ c^q\beta) \to (\beta^{-1}c^{p-q}\alpha,\ c^q\beta) \to (c^{p-q}\alpha\beta^{-1},\ c^q\beta)$$

and $G(\alpha, \beta) = F$ implies $G(\alpha\beta^{-1}, \beta) = F$. For $q < 0$

$$(c^p\alpha,\ c^q\beta) \to (c^p\alpha,\ \beta c^q) \to (\beta c^{p+q}\alpha,\ c^q\beta) \to (c^{p+q}\alpha\beta,\ c^q\beta)$$

and $G(\alpha, \beta) = F$ implies $G(\alpha\beta, \beta) = F$.

5. Some conjectures. It seems that the condition that $|r| + |s|$ is minimal is very powerful and of course we expect the following:

CONJECTURE 1. *If $N(r, s) = F$ and $|r| + |s|$ is minimal, then $|r| + |s| = 2$.*

This gives the main conjecture in the case in which we are working (F free on 2 generators).

The difficulty is that we don't know how many operations may be required to reduce $|r| + |s|$ when we know it may be reduced.

CONJECTURE 2. *If $|r| + |s|$ is not minimal, then at most three extended Nielsen operations are required to reduce it.*

In all examples we know if $(r, s) \xrightarrow{N} (a, b)$ then the only conjugations needed are cyclic conjugations.

CONJECTURE 3. *If $(r, s) \xrightarrow{N} (a, b)$, then only cyclic conjugations are required.*

Again there seems to be experimental evidence for the following:

CONJECTURE 4. *If $|r| + |s|$ is not minimal, then, considered as cyclic words, they have a common subword l such that*

$$|r| + |s| - 2|l| < \max\{|r|, |s|\}.$$

Finally, we remark that our main conjecture is equivalent to a kind of "sequential solvability." We say that (r, s) is a sequentially solvable pair if setting $r = 1 = s$ will yield

$$r' = 1 = s',\ r'' = 1 = s'',\ \cdots,\ a = 1 = b$$

with the stipulation that to get $r^{(k)} = 1 = s^{(k)}$ only $r^{(k-1)} = 1 = s^{(k-1)}$ are used. It is not hard to see that (r, s) is sequentially solvable if and only if $(r, s) \xrightarrow{N} (a, b)$.

This work was partially supported by NSF-GP626.

References

1. J. J. Andrews and M. L. Curtis, Free groups and handlebodies, to appear.
2. J. Nielsen, Om Regnig med ikke-kommutative Faktorer og dens Anvendelse i Gruppeteorien, Nordisk Mat. Tidskr., B (1921) 77–94.

SYSTEMS OF BOOLEAN EQUATIONS

RALPH M. TOMS, Renton, Washington

1. Introduction. In a solvable Boolean equation or system of equations each unknown has in general more than one value. Whitehead deduced necessary and sufficient conditions that a solution be unique [11]. A general proof of Whitehead's theorem has been given by Bernstein [1]. In a later paper, by Bernstein and Parker, the conditions are stated in several forms and an algorithm for determining the unique solution is developed [2]. This algorithm utilizes the disjunctive canonical form of a Boolean function. In this article a new uniqueness condition is developed and some new algorithms are presented.

2. Definitions and symbols.

DEFINITION 1. One of the simplest Boolean algebras is that defined on the set $\{0, 1\}$ with Boolean addition, $+$, (disjunction, inclusive union, logical sum, inclusive *OR*) and Boolean product, $\cdot$, (conjunction, intersection, logical product, *AND*) of any two members of the set defined by, $a+b=0$ if $a=b=0$; otherwise $a+b=1$ and $ab=1$ if $a=b=1$; otherwise $ab=0$. This particular Boolean algebra has been very useful in its applications to switching network theory [9, 10] and to the solution of a class of logical problems [3]. The remainder of this paper will be devoted exclusively to it.

DEFINITION 2. A variable x is a Boolean variable if and only if its range is the set $\{0, 1\}$. A letter is a symbol representing an independent Boolean variable. The dual x' of a letter is a Boolean variable such that $x'=0$ when $x=1$ and $x'=1$ when $x=0$. Both x and x' are called the literals associated with the letter x.

DEFINITION 3. A Boolean function is any expression which represents the combination of a finite set of symbols (each representing a constant or a variable) by the operations of $+$, $\cdot$, or $'$. Thus $a'b+0$ is a Boolean function provided that the symbols a and b are letters. The set of all Boolean functions of n independent Boolean variables will be called the finite Boolean algebra B_n. Two representations of elements in B_n are equivalent, $\equiv$, if and only if they have the same value for every set of values of the independent variables. The symbols $f(x_n)$ or $f(x_1, x_2, \cdots, x_n)$ will be used to represent members of B_n. The symbol, $=$, may be used for conditional equivalence. However, when there is no danger of confusion $f\equiv g$ may be written $f=g$.

DEFINITION 4. A monomial is a product of literals of B_n such that no letter occurs more than once in the product. A polynomial is the Boolean constant 1 or the sum of a set of monomials of B_n. If the set of monomials is empty the polynomial is the Boolean constant 0. An alterm is a sum of literals such that no letter occurs more than once in the sum. A conjunctive form of a Boolean function is a product of alterms.

DEFINITION 5. If C and D are monomials of a polynomial and x is an independent Boolean variable such that $A = Cx$ and $B = Dx'$ then CD is called the consensus of the monomials A and B.

DEFINITION 6. If a, b and c are variables such that $a = bc$ then $(a \leqq b, b \geqq a)$ and $(a \leqq c, c \geqq a)$. If A and B are monomials of a polynomial and $A \leqq B$ then A is said to subsume B.

3. Canonical forms. By Definition 3 the equivalence of two Boolean functions can be determined by evaluating each of them for every set of values of the independent variables. This is the familiar truth table method. Although this can be a useful method it is advantageous to be able to determine the equivalence between Boolean functions by a direct inspection of the representations of the functions.

There exist transformations such that every equivalent representation of a Boolean function is transformed into a unique representation of that function. These unique representations are called canonical forms. The two classical canonical forms are the disjunctive canonical form and the conjunctive canonical form [8]. Another useful canonical form is that due to Quine [6, 7] or Samson and Mills [9].

DEFINITION 7. The disjunctive canonical form of a Boolean function is given by,

$$f(x_n) = \sum_{i_1=0}^{1} \cdots \sum_{i_n=0}^{1} f(1^{i_1}, 1^{i_2}, \cdots, 1^{i_n}) \cdot (x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}),$$

where $x_j^i = x_j$ if $i = 0$; $x_j^i = x_j'$ if $i = 1$ and $(i, j = 1, 2, \cdots, n)$. Each term of the disjunctive canonical form is called a minterm. The coefficients

$$f(1^{i_1}, 1^{i_2}, \cdots, 1^{i_n})$$

will be called the discriminants of the disjunctive canonical form. In what follows the symbol $Q(x_n)$ or $Q(x_1, x_2, \cdots, x_n)$ will represent the disjunctive canonical form of a given Boolean function $f(x_n)$.

The following well-known properties of B_n will be utilized (for proofs see [8]).

PROPERTY 1. For any Boolean function $f(x_1, x_2, \cdots, x_n)$ in B_n,

$$f(x_1, x_2, \cdots, x_n) = x_n f(x_1, x_2, \cdots, x_{n-1}, 1) + x_n' f(x_1, x_2, \cdots, x_{n-1}, 0).$$

This relation is sometimes referred to as the fundamental theorem of Boolean algebra. Thus $Q(x_n)$ of $f(x_n)$ may be written,

$$Q(x_n) = x_n Q(x_1, x_2, \cdots, x_{n-1}, 1) + x_n' Q(x_1, x_2, \cdots, x_{n-1}, 0).$$

It should be noted that the polynomials $Q(x_1, x_2, \cdots, x_{n-1}, 1)$ and $Q(x_1, x_2, \cdots, x_{n-1}, 0)$ are disjunctive canonical forms of Boolean functions in $n-1$ independent variables.

PROPERTY 2. If a, b and c are Boolean variables then $a+b+bc=a+b$. This procedure is called a deletion. The performance of all possible deletions on a Boolean function $f(x_n)$ is called a deletion iteration on $f(x_n)$.

PROPERTY 3. The disjunctive canonical form of the dual of a Boolean function is the sum of the minterms of the set complement of the disjunctive canonical form of the given Boolean function relative to the set of 2^n minterms of B_n.

DEFINITION 8. The following transformation, discovered by Samson and Mills [9], will reduce any member of B_n to Quine's canonical form.

(i) Transform the given Boolean function $f(x_n)$ into an equivalent polynomial representation $P(x_n)$.

(ii) If A and B are monomials of $P(x_n)$ such that $A \leqq B$ then remove B from the representation $P(x_n)$.

(iii) If xA and $x'B$ are monomials of $P(x_n)$ such that the consensus AB is not the Boolean constant 0 and does not subsume any monomial of $P(x_n)$ then add the consensus AB to $P(x_n)$.

(iv) If the monomials A and A' occur in the representation then the canonical form is the Boolean constant 1.

(v) When a consensus is added in step (iii) then (ii), (iii), and (iv) are repeated until no new consensus can be added.

Proofs that this representation is canonical are given by Quine [6, 7] and Laxdal [5]. Digital computer programs have been developed by Laxdal [5] and Witcraft [12] which reduce Boolean functions to Quine's canonical form. The monomials of Quine's canonical form are called prime implicants. The symbol $QF\colon f(x_n)$ will be read, "Quine's canonical form of $f(x_n)$."

The following useful theorem is due to Ghazala [4]:

THEOREM A. *If $f(x_n)$ is a Boolean function in conjunctive form and if*

(i) *all the indicated multiplications are performed,*

(ii) *all products xx' are dropped from the resulting polynomial $P(x_n)$,*

(iii) *a deletion iteration is performed on $P(x_n)$,*

then the resulting polynomial is $QF\colon f(x_n)$.

4. Systems of Boolean Equations. Consider the system of m Boolean equations in n unknowns,

$$f_1(x_n) = g_1(x_n), f_2(x_n) = g_2(x_n), \cdots, f_m(x_n) = g_m(x_n). \tag{1}$$

Any Boolean equation $f=g$ can be written in the form $fg'+f'g=0$ and hence, the system (1) becomes

$$\begin{aligned} f_1(x_n)g_1'(x_n) + f_1'(x_n)g_1(x_n) &= 0 \\ f_2(x_n)g_2'(x_n) + f_2'(x_n)g_2(x_n) &= 0 \\ &\cdots \\ f_m(x_n)g_m'(x_n) + f_m'(x_n)g_m(x_n) &= 0. \end{aligned} \tag{2}$$

If a solution exists for the system (2) the left hand side of every equation of the system must have the value 0 for some set of values of the variables. The system (2) can therefore be reduced to the single Boolean equation

$$H(x_n) = \sum_{i=1}^{m} f_i(x_n)g_i'(x_n) + f_i'(x_n)g_i(x_n), \tag{3}$$

where $H(x_n)$ is a function of n independent variables. Any solution to the equation (3) is a solution to the system (2) and therefore is a solution of system (1). Thus it is sufficient to examine the solution of the single Boolean equation (3). A solution to a Boolean equation such as (3) will be written as ordered n-tuples of the Boolean constants 0 and 1. For example (0, 1) is a solution of $x+y'=0$.

Whitehead's conditions for the uniqueness of a solution to the equation (3) require the use of the disjunctive canonical form $Q(x_n)$ of $H(x_n)$ [11]. That is,

THEOREM B. *The necessary and sufficient conditions that the equation $Q(x_n)=0$ have a unique solution are*

(i) $a_1a_2 \cdots a_2n=0$ *and* (ii) $a_i'a_j'=0$ $(i\neq j)$,

where the a_i are the discriminants of $Q(x_n)$.

It should be noted that this theorem implies that if $Q(x_n)=0$ has a unique solution exactly one discriminant of $Q(x_n)$ has the value 0.

THEOREM 1. *If $T=\{a_j \mid a_j=1\}$, where the a_j are the discriminants of $Q(x_n)$ and t is the cardinality of T, then the number of solutions to the equation $Q(x_n)=0$ is 2^n-t.*

Proof. The proof is by induction on n. For the case $n=1$ the proof is trivial. Assume now that the statement is true for $n=k$ and consider the equation $Q(x_{k+1})=0$. By the fundamental theorem (Property 1):

$$Q(x_{k+1}) = x_{k+1}Q(x_1, x_2, \cdots, x_k, 1) + x_{k+1}'Q(x_1, x_2, \cdots, x_k, 0).$$

Let the cardinality of T for $Q(x_1, x_2, \cdots, x_k, 1)$ be t_1 and for $Q(x_1, x_2, \cdots, x_k, 0)$ let it be t_2. For the case where $x_{k+1}=1$ it is clear that $Q(x_{k+1})$ has the value 0 only if $Q(x_1, x_2, \cdots, x_k, 1)$ has the value 0. By the induction hypothesis, this implies that $Q(x_{k+1})$ has 2^k-t_1 solutions for $x_{k+1}=1$. By a symmetrical argument, if $x_{k+1}'=1$ then $Q(x_{k+1})=0$ has 2^k-t_2 solutions. Since these cases are mutually exclusive it follows that there are $2^k-t_1+2^k-t_2=2^{k+1}-(t_1+t_2)$ solutions of $Q(x_{k+1})=0$. Clearly the set of discriminants of $Q(x_{k+1})$ is the union of the sets of discriminants of $Q(x_1, x_2, \cdots x_k, 1)$ and $Q(x_1, x_2, \cdots, x_k, 0)$. Therefore the cardinality of T for $Q(x_{k+1})$ is t_1+t_2. Hence, by mathematical induction, the theorem is true for all n.

In addition to indicating the number of solutions, it should also be observed that Theorem 1 yields all of the solutions to a given equation. For example, consider the equation

$$Q(x, y, z) = x'yz + xy'z + x'y'z + x'y'z' + xy'z' + xyz = 0.$$

Since $Q(x, y, z)$ is in disjunctive canonical form, by Theorem 1 there are exactly two solutions to the equation. By taking the set complement of the set of minterms of $Q(x, y, z)$ and setting each of these minterms equal (equivalent) to 1 the solutions are readily obtained. That is,

$$xyz' = 1 \text{ for } (1, 1, 0), \qquad x'yz' = 1 \text{ for } (0, 1, 0),$$

and these are the two solutions to the given equation.

Theorem 1 is simply a generalization of Theorem B and is, of course, no more than a formalized truth table. Thus the utilization of Theorem 1 necessitates the examination of 2^n monomials. For large n this may not be practical except in particular instances where the set complement of the set of minterms is small and readily obtained. The following theorem leads to methods that may have a wider range of applicability.

THEOREM 2. *A necessary and sufficient condition that the equation $f(x_n)=0$ have a unique solution is that $QF{:}f(x_n)$ is an alterm in n letters.*

Proof. The condition is sufficient for if,

$$QF{:}f(x_n) = \sum_{i=1}^{n} \alpha_i = 0,$$

where the α_i are literals, then every α_i must be exactly 0 and a unique solution is determined.

If there is a unique solution to the equation then by Theorem B (or Theorem 1) exactly one discriminant of the disjunctive canonical form of $f(x_n)$ is 0. If M is the minterm whose discriminant is 0 then by Property 3,

$$M = (Q(x_n))'.$$

Now,

$$f(x_n) = Q(x_n) = (Q'(x_n))' = \sum_{i=1}^{n} \alpha_i,$$

where $\sum_{i=1}^{n} \alpha_i$ is an alterm. The term $\sum_{i=1}^{n} \alpha_i$ is a conjunctive form (a conjunction of one alterm). By Theorem A the alterm is in $QF{:}f(x_n)$ and this shows the necessity of the condition.

The theorem just presented can be used to obtain the unique solution (if it exists) to a given equation $f(x_n)=0$. By the transformation shown in Definition 8 the function $f(x_n)$ is transformed into Quine's canonical form and the solution is easily obtained. As an example of this procedure consider the equation,

$$f(w, x, y, z) = xz + w'x'z + x'yz + xz' + xyz' + w'x'z' + x'z' = 0.$$

By the transformation of Definition 8

$$QF{:}f(w, x, y, z) = w' + x + y + z' = 0$$

and clearly the solution is (1, 0, 0, 1).

THEOREM 3. *If $QF{:}\,f(x_n)$ has in its representation prime implicants which are literals then in any solution to $f(x_n)=0$ the values of the variables associated with the literals are uniquely determined.*

Proof. Suppose that $QF{:}f(x_n)=P+\psi_i$, where P is a polynomial no monomial of which is a literal and ψ_i is an alterm. Then if $f(x_n)$ has the value 0 for a set of values of the independent variables every literal of $\psi_i\equiv 0$. In that case the variable associated with a particular literal is uniquely determined.

THEOREM 4. *If $QF{:}\,f(x_n)$ is not the Boolean constant 1 and if a letter appearing in the representation of $f(x_n)$ does not appear in the $QF{:}\,f(x_n)$ then the variable associated with the missing letter is arbitrary in any solution of $f(x_n)=0$.*

Proof. It is evident that a letter can be deleted only by a multiplication by 0, a summation with its dual to yield 1 or by a deletion iteration. That is, suppose that x is a letter appearing in $f(x_n)$ but not in $QF{:}f(x_n)$ then if A and B are not functions of x:

(i) $f(x_n)$ has the form $A+0\cdot x$ and the value of x is arbitrary, or

(ii) $f(x_n)$ has the form $A+B(x+x')$ and the value of x is arbitrary, or

(iii) the form $Ax+A$ appears in the representation of $f(x_n)$ during the Quine reduction and by Property 2 the value of x is arbitrary.

THEOREM 5. *If $f(x_n)$ is a Boolean polynomial, then a necessary and sufficient condition that the equation $f(x_n)=0$ will have a unique solution is that $QF{:}\,(f(x_n))'$ is a monomial in n literals.*

Proof. This is the dual of Theorem 3.

The above results will yield several algorithms for determining unique solutions to systems of Boolean equations. Often they will determine arbitrary variables and yield complete sets of solutions. For example, consider the following algorithm.

Given the equation $f(x_n)=0$;

1. transform $f(x_n)$ to an equivalent polynomial $P(x_n)$,
2. perform a deletion iteration,
3. determine the conjunctive form $(P(x_n))'$,
4. perform a deletion iteration,
5. perform the indicated multiplications,
6. perform a deletion iteration.

If the above procedure results in a monomial representation $M(x_n)$ of $f(x_n)$ in n literals the equation has a unique solution which is obtained by solving the

equation $M(x_n)=1$. As an example consider the equation

$$f(x, y) = y(x + x'y) + xy'$$

then,

$$\begin{aligned} f(x, y) &= xy + x'y + xy' \\ f(x, y)' &= (x' + y')(x + y')(x' + y) \\ f(x, y)' &= x'y' \\ f(x, y) &= x + y = 0 \end{aligned}$$

and the unique solution is (0, 0). As another example consider the equation

$$f(x, y, z) = x(x + yz) + yx' = 0$$

then,

$$\begin{aligned} f(x, y, z) &= x + xyz + yx' \\ f(x, y, z) &= x + yx' \\ f(x, y, z)' &= x'(y' + x) = x'y' \\ f(x, y, z) &= x + y = 0 \end{aligned}$$

and the solutions are (0, 0, 0) and (0, 0, 1) since z is arbitrary.

The primary results of this paper were taken from a thesis submitted to Oregon State University in partial fulfillment of requirements for the M.S. degree. The author is indebted to Professor Harry E. Goheen for his able guidance during this research.

References

1. B. A. Bernstein, Note on the condition that a Boolean equation have a unique solution, Amer. J. Math., 54 (1932) 417–418.

2. ——— and W. L. Parker, On uniquely solvable Boolean equations, Univ. Calif. Pub. Math., 3 (1955) 1–30.

3. T. J. Fletcher, The solution of inferential problems by Boole algebra, Math. Gaz., 35 (1952) 183–188.

4. M. J. Ghazala, Irredundant disjunctive and conjunctive forms of a Boolean function, IBM J. Res. Dev., 1 (1957) 171–176.

5. A. L. Laxdal, A mechanization of Quine's canonical form, Masters Thesis, Oregon St. Univ., 1959.

6. W. V. Quine, The problem of simplifying truth functions, this MONTHLY, 59 (1952) 521–530.

7. ———, A way to simplify truth functions, this MONTHLY, 62 (1955) 627–630.

8. P. Rosenbloom, The elements of mathematical logic, Dover, New York, 1950.

9. E. W. Samson and B. E. Mills, Circuit minimization algebra and algorithms for new Boolean canonical expressions, U. S. A. F. Cam. Res. Cent. Tech. Rep. No. 54-21 (1954).

10. C. Shannon, A symbolic analysis of relay and switching circuits, Trans. Amer. Inst. Elec. Engrs., 57 (1938) 713–723.

11. A. N. Whitehead, Memoir on the algebra of symbolic logic, Amer. J. Math., 23 (1901) 139–165; 297–316.

12. D. A. Witcraft, The mechanization of logic II, Masters Thesis, Oregon St., Univ., 1960.

THE CONSTRUCTION OF PROJECTIVE PLANES FROM GENERALIZED TERNARY RINGS

J. R. WESSON, Vanderbilt University

1. Introduction. It is well known that any projective plane can be coordinatized from a ternary ring with a zero and a unit [**1**, p. 118; **2**, p. 3; **4**, p. 247; **5**, p. 353]. In this paper a ternary ring R is defined without the usual assumptions of the existence of a zero and a unit, and it is shown that R induces a projective plane. From the ternary ring R there is constructed a ternary ring R' with a zero and a unit. The planes induced by R and R' are isomorphic.

2. Construction of projective planes from ternary rings.

DEFINITION 2.1. *A ternary ring is a set* $R=\{a, b \cdots\}$ *of at least two elements with a ternary operation* xyz (*that is, typified by* xyz) *such that*:

T1. *For all* $x, y, z \in R$, xyz *is a unique element of* R.

T2. *If* $a, b, c \in R$, *there exists a unique* $z \in R$ *such that* $abz=c$.

T3. *If* $a, b, c, d \in R$ *with* $a \neq c$, *there exists a unique* $x \in R$ *such that* $xab=xcd$.

T4. *If* $a, b, c, d \in R$ *with* $a \neq c$, *there exists a unique pair* $x, y \in R$ *such that* $axy=b$, $cxy=d$.

DEFINITION 2.2. *A projective plane* (*or plane*) *is a set S of "points" with certain subsets called "lines" such that*:

P1. *Any two distinct points are contained in exactly one line.*

P2. *Any two distinct lines contain exactly one common point.*

P3. *S contains at least four points, no three of which are on the same line.*

We now construct a plane from a ternary ring.

THEOREM 2.1. *Let* R *be a ternary ring with operation* xyz. *Take as points* (∞), (m), (a, b) *where* $a, b, m \in R$ *and* $\infty \notin R$. *Choose as lines* L_∞, $x=a$, $y=xmb$, *where* $a, b, m \in R$. *Define the following* (*and no other*) *incidences of points and lines*:

$$(\infty) \in L_\infty, \qquad (m) \in y = xmb,$$
$$(\infty) \in x = a, \qquad (a, b) \in x = a,$$
$$(m) \in L_\infty, \qquad (a, b) \in y = xmc \text{ if and only if } amc = b.$$

Then the points, lines, and incidences so defined form a projective plane π, [**5**, p. 355].

Proof. First observe that the point (∞) cannot lie on a line $y=xmb$. Similarly, $(m) \notin x=a$ and $(a, b) \notin L_\infty$.

Next, we prove that any two distinct points lie on exactly one line. Only the more difficult of five cases are considered. The points (m) and (a, b) lie on the line $y=xmc$, where c is determined uniquely by $amc=b$ according to T2. The distinct points (a_1, b_1), (a_2, b_2) lead to two cases. If $a_1=a_2$, the points lie on

$x=a_1$, but they do not lie on a line $y=xmc$ for this would imply $a_1mc=a_2mc=b_1=b_2$, a contradiction to the distinctness of the points. On the other hand, if $a_1\neq a_2$, the points (a_1, b_1), (a_2, b_2) lie on $y=xmc$, where the equations $a_1mc=b_1$, $a_2mc=b_2$ determine m, c uniquely by T4. Hence P1 holds.

To establish P2 it is necessary only to show that any two distinct lines have a point in common, for the intersection of two distinct lines in more than one point is prevented by P1. Again, only two of five cases are considered. The lines $x=a$, $y=xmb$ contain (a, amb). The distinct lines $y=xm_1b_1$, $y=xm_2b_2$ contain (m_1), if $m_1=m_2$. On the other hand, if $m_1\neq m_2$, the lines contain (u, v), where u is determined uniquely by $um_1b_1=um_2b_2$ (T3) and v is determined uniquely by $v=um_1b_1$ (T1).

Finally, four points, no three collinear, are displayed. Consider the noncollinear points (∞), (a, a), (b, b), with $a\neq b$. Let (u) be a point not on the line joining (a, a) and (b, b). Then the four points (∞), (a, a), (b, b), (u) are the ones required by P3.

This completes the proof of the theorem. The plane π is called the *plane induced by R.*

3. Introduction of a zero in a ternary ring. The introduction of a zero in a ternary ring R is made in two steps. First a "left" zero is introduced to change R into another ternary ring R_1. Then a zero is introduced in R_1 to form a ternary ring R_2. The planes induced by R and R_2 are isomorphic, and each is the dual of the plane induced by R_1.

DEFINITION 3.1. *Two planes π_1 and π_2 are isomorphic* [5, p. 348] *if there exists a one-one mapping α of the points of π_1 onto the points of π_2 and a one-one mapping β of the lines of π_1 onto the lines of π_2 such that $P\in l$ in π_1 implies $P\alpha\in l\beta$ in π_2. (P represents a point in π_1, l represents a line of π_1.)*

DEFINITION 3.2. *Two planes π_1 and π_2 are dual* [5, p. 347] *if there exists a one-one mapping α of the points of π_1 onto the lines of π_2 and a one-one mapping β of the lines of π_1 onto the points of π_2 such that $P\in l$ in π_1 implies $l\beta\in P\alpha$ in π_2.*

Isomorphism of planes is an equivalence relation. Duality of planes is a symmetric relation, and planes dual to the same plane are isomorphic.

THEOREM 3.1. *Let R, xyz be a ternary ring (that is, let R be a ternary ring with operation xyz). Define a new ternary operation $\langle xyz\rangle$ by $yx\langle xyz\rangle=y0z$, where 0 is any element of R. Then R, $\langle xyz\rangle$ is a ternary ring satisfying $\langle 0yz\rangle=z$ (for all y, z), and the plane induced by R, $\langle xyz\rangle$ is dual to the plane induced by R, xyz.*

Proof. First it is proved that $\langle xyz\rangle$ satisfies T1 through T4. We have, of course, that xyz satisfies T1 through T4.

T2 for the operation xyz implies T1 for $\langle xyz\rangle$.

$\langle abz\rangle=c$ if and only if $ba\langle abz\rangle=bac$, or if and only if $b0z=bac$. Hence there is a unique z satisfying $\langle abz\rangle=c$.

Let $a \neq c$. There exists a unique x such that $\langle xab \rangle = \langle xcd \rangle$ if and only if there exists a unique pair x, u such that $\langle xab \rangle = u$, $\langle xcd \rangle = u$. This system is equivalent to $ax\langle xab \rangle = axu$, $cx\langle xcd \rangle = cxu$ or $axu = a0b$, $cxu = c0d$. Therefore, the existence of a unique pair x, u is assured by T4.

Again, let $a \neq c$. The system $\langle axy \rangle = b$, $\langle cxy \rangle = d$ has a unique solution x, y if and only if the system $xab = x0y$, $xcd = x0y$ has a unique solution x, y. Now x is determined uniquely by $xab = xcd$ according to T3. Then y is determined uniquely by $x0y = xab$ and T2.

Therefore, R, $\langle xyz \rangle$ is a ternary ring. Also, the definition of $\langle xyz \rangle$ gives $y0\langle 0yz \rangle = y0z$. Hence T2 implies $\langle 0yz \rangle = z$.

Next we show that the two planes induced by the ternary rings are dual. Let π_1 be the plane induced by R, xyz and let π_2 be the plane induced by R, $\langle xyz \rangle$. The following correspondence gives the duality.

π_1	π_2	π_1	π_2
(∞)	$\leftrightarrow L_\infty$	L_∞	$\leftrightarrow (\infty)$
(m)	$\leftrightarrow x = m$	$x = a$	$\leftrightarrow (a)$
$(a, a0b)$	$\leftrightarrow y = \langle xab \rangle$	$y = xmc$	$\leftrightarrow (m, c)$

We need only to verify that if in π_1 the point P belongs to the line l, then in π_2 the image of l belongs to the image of P. Only one of six cases is considered. If $(a, a0b) \in y = xmc$ in π_1, then $a0b = amc$, $am\langle mab \rangle = amc$, $\langle mab \rangle = c$, and $(m, c) \in y = \langle xab \rangle$ in π_2. This completes the proof of the theorem.

THEOREM 3.2. *Let R, xyz be a ternary ring satisfying the identity $0yz = z$ for a fixed element 0. Define a new operation $\langle xyz \rangle$ by $yx\langle xyz \rangle = y0z$. Then R, $\langle xyz \rangle$ is a ternary ring with zero; that is, $\langle x0z \rangle = \langle 0yz \rangle = z$ for all x, y, z.*

Proof. That R, $\langle xyz \rangle$ is a ternary ring satisfying $\langle 0yz \rangle = z$ follows from Theorem 3.1. Now $0x\langle x0z \rangle = 00z$ from the definition of $\langle xyz \rangle$. Also, $0x\langle x0z \rangle = \langle x0z \rangle$ by hypothesis. Hence $\langle x0z \rangle = 00z = z$. This completes the proof.

THEOREM 3.3. *Let R, xyz be a ternary ring and let 0 be a fixed element of R. Define $\langle xyz \rangle$ by $yx\ \langle xyz \rangle = y0z$, and define $[xyz]$ by $\langle yx\ [xyz] \rangle = \langle y0z \rangle$. Then R, $[xyz]$ is a ternary ring satisfying $[x0z] = [0yz] = z$, and the plane induced by R, $[xyz]$ is isomorphic to the plane induced by R, xyz.*

Proof. It is necessary only to refer to Theorems 3.1 and 3.2 and to recall that the dual of the dual of a plane π is π itself.

For a ternary ring R with a zero (an element 0 such that $x0z = 0yz = z$) the following properties are immediate.

Z1. *If a, b, $c \in R$ with $a \neq 0$, there exists a unique element $x \in R$ such that $xab = c$, and there exists a unique element $y \in R$ such that $ayb = c$.*

Z2. $ab0 = 0$ *if and only if either* $a = 0$ *or* $b = 0$.

Z3. *If* $a \neq 0$ *and* $abc = adc$, *then* $b = d$.
Z4. *If* $a \neq 0$ *and* $bad = cad$, *then* $b = c$.

Z1 is proved by taking $c = 0$ in T3 and by taking $c = 0$ in T4. Then Z2, Z3 and Z4 are consequences of Z1.

4. Introduction of a unit [3, pp. 8, 56]. For a ternary ring R with a zero, left and right divisions will be defined. Then a new operation on the elements of R gives a ternary ring with zero and unit. The plane induced by the new ternary ring is isomorphic to the original plane.

DEFINITION 4.1. *Let* R, xyz *be a ternary ring with zero. For any nonzero element* $\alpha \in R$ *define* (i) $x/\alpha = y$ *if and only if* $x = y\alpha 0$. (ii) $\alpha\backslash x = y$ *if and only if* $x = \alpha y 0$.

Property Z1 assures the unique existence of x/α and $\beta\backslash x$. Similarly, the three statements $a = b$, $a/\alpha = b/\alpha$, $\beta\backslash a = \beta\backslash b$ are equivalent. That $0/\alpha = \beta\backslash 0 = 0$ is obvious, and it is not difficult to verify that $(x/\alpha)\alpha 0 = x$, $\beta(\beta\backslash y)0 = y$, $x\alpha 0/\alpha = x$, and $\beta\backslash\beta y 0 = y$.

THEOREM 4.1. *Let* R, xyz *be a ternary ring with zero* (0), *and define* $\langle xyz\rangle$ *by* $\langle xyz\rangle = (x/\alpha)(\beta\backslash y)z$, *where* α, β *are fixed nonzero elements of* R. *Then* R, $\langle xyz\rangle$ *is a ternary ring with zero* (0) *and with unit* $e = \beta\alpha 0$ (*that is,* $ex0 = xe0 = x$ *for all* x). *Furthermore, the plane induced by* R, $\langle xyz\rangle$ *is isomorphic to the plane induced by* R, xyz.

Proof. The operation $\langle xyz\rangle$ is uniquely defined. The equation $\langle abz\rangle = c$ is equivalent to $(a/\alpha)(\beta\backslash b)z = c$, and therefore T2 holds. Next, $\langle xab\rangle = \langle xcd\rangle$ if and only if $(x/\alpha)(\beta\backslash a)b = (x/\alpha)(\beta\backslash c)d$, and hence for $a \neq c$, x/α and x are uniquely determined. This establishes T3. Also, $\langle axy\rangle = b$, $\langle cxy\rangle = d$ if and only if $(a/\alpha)(\beta\backslash x)y = b$, $(c/\alpha)(\beta\backslash x)y = d$, and for $a \neq c$, x and y are uniquely determined. This establishes T4. Clearly $\langle x0z\rangle = \langle 0yz\rangle = z$. Furthermore, the element $e = \beta\alpha 0$ is a unit, for $xe0 = (x/\alpha)(\beta\backslash e)0 = (x/\alpha)\alpha 0 = x$ and $ey0 = (e/\alpha)(\beta\backslash y)0 = \beta(\beta\backslash y)0 = y$. Hence R, $\langle xyz\rangle$ is a ternary ring with zero and unit.

Let π_1 be the plane induced by R, xyz and let π_2 be the plane induced by R, $\langle xyz\rangle$. The following correspondence is an isomorphism between π_1 and π_2.

π_2		π_1	π_2		π_1
(∞)	$\leftrightarrow$	(∞)	L_∞	$\leftrightarrow$	L_∞
(m)	$\leftrightarrow$	$(\beta\backslash m)$	$x = a$	$\leftrightarrow$	$x = a/\alpha$
(a, b)	$\leftrightarrow$	$(a/\alpha, b)$	$y = \langle xmb\rangle$	$\leftrightarrow$	$y = x(\beta\backslash m)b$

It is not difficult to show that the correspondence preserves incidence between points and lines. For example, let $(a, b) \in y = \langle xmc\rangle$ in π_2. Then $b = \langle amc\rangle = (a/\alpha)(\beta\backslash m)c$ and $(a/\alpha, b) \in y = x(\beta\backslash m)c$ in π_1.

Presented March 20, 1964 to the Southeastern Section of the Association at Charleston, South Carolina.

References

1. L. M. Blumenthal, A modern view of geometry, Freeman, San Francisco and London, 1961.
2. R. H. Bruck, Recent advances in the foundations of Euclidean plane geometry, this MONTHLY, 62, part II (1955) 2–17.
3. ———, A survey of binary systems, Springer, Berlin-Göttingen-Heidelberg, 1958.
4. M. Hall, Projective planes, Trans. Amer. Math. Soc., 54 (1943) 229–277.
5. ———, The theory of groups, Macmillan, New York, 1959.

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

Material for this department should be sent to J. H. Curtiss, University of Miami, Coral Gables, Florida 33146.

BOOLEAN-LIKE FUNCTIONS

E. D. GOODRICH, Ohio University

Introduction. In [1] Foster introduced the concept of a Boolean-like ring with the following definition:

A ring R is Boolean-like if and only if

(i) It is commutative, with unit element.

(ii) It is of characteristic 2.

(iii) Each element x in R may be expressed uniquely as $x=x_i+x_n$ where x_i and x_n are idempotent and nilpotent elements, respectively, of R.

(iv) $nn^*=0$ for all nilpotent elements n, n^*.

In addition to the two binary operations of R we define a unary operation as $a'=1+a$, where 1 is the unit element. It was shown also in [1] that every Boolean ring is Boolean-like but not conversely. In one respect, a Boolean-like ring differs from a Boolean ring in the sense that not every element need be idempotent. It can easily be shown that if 0 is the only nilpotent element of R, then R must be a Boolean ring.

In [2] Hohn defines a Boolean function of n variables to be of the form

$$f(x_1, x_2, \cdots, x_n) = \sum_{(e)} f(e_1, e_2, \cdots, e_n)x_1^{e_1}x_2^{e_2} \cdots x_n^{e_n},$$

where the summation extends over all combinations (e) of n zeros and ones and $x_j^0=x_j'$ $x_j^1,=x_j$ where $'$ is the unary operation of the Boolean ring. In particular, a function of one variable is of the form $f(x)=f(1)x+f(0)x'$; a function of two variables is of the form $f(x, y)=f(1, 1)xy+f(0, 1)x'y+f(1, 0)xy'+f(0, 0)x'y'$.

It is the purpose of this paper to define a Boolean-like function of n variables and to show that all functions of one, two and n variables are of a form similar to that of Boolean functions.

DEFINITION 1. *Let $x_1, x_2, \cdots, x_n$ be variables whose domain is a Boolean-like ring B. Then a function $f(x_1, x_2, \cdots, x_n)$, built up from these variables and from elements of B by a finite number of applications of the binary and unary operations of B is called a Boolean-like function.*

THEOREM 1. *Any Boolean-like function of one variable is of the form $f(x)=ax_i+bx_i'+cx_ix_n+dx_i'x_n$ where x_i and x_n are the idempotent and nilpotent parts, respectively, of x and a, b, c, d are elements of a Boolean-like ring B.*

Proof. It is immediate that the functions c, x_i, x_i' and x_n are of the required form since

$$\begin{aligned} c &= cx_i + cx_i' + 0x_ix_n + 0x_i'x_n \quad \text{(for } c \text{ any element of } B) \\ x_i &= 1x_i + 0x_i' + 0x_ix_n + 0x_i'x_n \\ x_i' &= 0x_i + 1x_i' + 0x_ix_n + 0x_i'x_n \\ x_n &= 0x_i + 0x_i' + 1x_ix_n + 1x_i'x_n \end{aligned}$$

Now if $f(x)=ax_i+bx_i'+cx_ix_n+dx_i'x_n$ and $g(x)=ex_i+fx_i'+gx_ix_n+hx_i'x_n$ are Boolean-like functions then $f(x)+g(x)$, $f(x)g(x)$ and $(f(x))'$ are also of the form since

$$\begin{aligned} f(x)+g(x) &= (a+e)x_i + (b+f)x_i' + (c+g)x_ix_n + (d+h)x_i'x_n \\ f(x)g(x) &= (ae)x_i + (bf)x_i' + (ag+ec)x_ix_n + (bh+fd)x_i'x_n \\ (f(x))' &= 1 + (ax_i + b\,x_i' + cx_ix_n + dx_i'x_n) \\ &= (x_i + x_i') + (ax_i + bx_i' + cx_ix_n + dx_i'x_n) \\ &= (1+a)x_i + (1+b)x_i' + cx_ix_n + dx_i'x_n \\ &= a'x_i + b'x_i' + cx_ix_n + dx_i'x_n. \end{aligned}$$

Since every Boolean-like function is built up from the variables and elements of B with a finite number of applications of the operations, then it follows that every function of one variable is of the stated form.

By the definition of a Boolean-like ring, every element b can be represented uniquely as $b=b_i+b_n$; if $c, d \in B$ then $c=c_i+c_n$ and $d=d_i+d_n$. Hence

$$cx_ix_n = (c_i + c_n)x_ix_n = c_ix_ix_n + c_nx_ix_n = c_ix_ix_n \quad \text{and} \quad dx_i'x_n = d_ix_i'x_n.$$

Therefore, without loss of generality, one need consider only idempotent elements as coefficients of the x_ix_n and $x_i'x_n$ terms since the nilpotent part adds nothing to the function. Also since $1=1+0$ and $0=0+0$ then $f(1)=a$ and $f(0)=b$. Now we define a Boolean-like function of one variable to be of the form $f(x)=f(1)x_i+f(0)x_i'+cx_ix_n+dx_i'x_n$ where c and d are idempotent elements of B.

The question naturally arises as to the uniqueness of a function of one variable, that is, if two functions are equal, what can be said concerning the coefficients of like terms? This will be resolved with the aid of the following theorem.

THEOREM 2. *Let B be a Boolean-like ring and $Q=\{x: x^2=x$ and $xn=0$ for all n such that $n^2=0\}$. Then Q is an ideal in B.*

Proof. Obviously $0\in Q$. Let $a, b\in Q$. Then if n is any nilpotent element, $(a+b)n=an+bn=0+0=0$. Thus $a+b\in Q$. Also for $a\in B$, $b\in Q$, $(ab)n=a(bn)=a0=0$. Hence $ab\in Q$ and Q is an ideal. Q will be called the "0-like ideal" of B. Now to return to the question of uniqueness.

THEOREM 3. *Let $f(x)=f(1)x_i+f(0)x_i'+cx_ix_n+dx_i'x_n$ and $g(x)=g(1)x_i+g(0)x_i'+\bar{c}x_ix_n+\bar{d}x_i'x_n$ and let Q be the 0-like ideal. Then $f(x)=g(x)$ if and only if $f(1)=g(1)$, $f(0)=g(0)$, $c\equiv\bar{c}$(mod Q) and $d\equiv\bar{d}$ (mod Q), where $m\equiv n$ (mod Q) means $m+n\in Q$.*

Proof. Assume that $f(x)=g(x)$. Now $f(1)=g(1)$ and $f(0)=g(0)$. Consider now z_n to be any nilpotent element, that is $z_n=0+z_n$. Hence $f(z_n)=f(0)+dz_n=g(0)+\bar{d}z_n$. Thus $dz_n=\bar{d}z_n$ which implies that $(d+\bar{d})z_n=0$. Since d and $\bar{d}$ are idempotent then $d+\bar{d}$ is idempotent. Hence $d+\bar{d}\in Q$. Thus $d+\bar{d}\equiv 0$ (mod Q) which implies that $d\equiv\bar{d}$ (mod Q). To verify that $c\equiv\bar{c}$ (mod Q) let z be any unipotent element, that is $z=1+z_n$, and evaluate $f(z)$.

Assume now that the conditions hold. Hence $c+\bar{c}\equiv 0$ (mod Q) and $d+\bar{d}\equiv 0$ (mod Q). For any nilpotent element x_n, $(c+\bar{c})x_n=0$ and $(d+\bar{d})x_n=0$. Thus $cx_n=\bar{c}x_n$ and $dx_n=\bar{d}x_n$. Also for x_i any idempotent element, $cx_ix_n=\bar{c}x_ix_n$ and $dx_i'x_n=\bar{d}x_i'x_n$. Hence $f(x)=g(x)$.

We may conclude that functions of one variable are unique up to equivalence classes of B modulo the 0-like ideal. If all elements of B are idempotent and 0 is the only nilpotent element (all idempotent elements force 0 to be the only nilpotent element) then $B=Q$ and $f(x)=f(1)x+f(0)x'$.

THEOREM 4. *Any Boolean-like function of two variables is of the form*

$$\begin{aligned} f(x, y) = f(1, 1)x_iy_i &+ f(0, 1)x_i'y_i + f(1, 0)x_iy_i' + f(0, 0)x_i'y_i' \\ &+ (a_0x_iy_i + a_1x_i'y_i + a_2x_iy_i' + a_3x_i'y_i')y_n \\ &+ (a_4x_iy_i + a_5x_i'y_i + a_6x_iy_i' + a_7x_i'y_i')x_n, \end{aligned}$$

where a_j $(j=0, 1, 2, 3, \cdots, 7)$ are idempotent elements of a Boolean-like ring.

Proof. The proof follows the same outline as Theorem 1. Again if the ring happens to be Boolean then $f(x, y)=f(1, 1)xy+f(0, 1)x'y+f(1, 0)xy'+f(0, 0)x'y'$ which is the Boolean function of two variables.

THEOREM 5. *Let $f(x, y)$ and $g(x, y)$ be Boolean-like functions. Then $f(x, y)=g(x, y)$ if and only if coefficients of like terms are congruent modulo the 0-like ideal.*

Proof. The proof is similar to that of Theorem 3.

A function of two variables, say $f(x, y)$, where the domain of the variables is a Boolean-like ring B is actually a function which relates to each pair of ele-

ments x and y of B an element of B (that is, $f(x, y)$ is a binary operation defined on B). One may consider the set of all functions of two variables as a set of binary operations defined on B. With this in mind, the following theorem will be proved.

THEOREM 6. *A necessary and sufficient condition that a binary operation* $\circ$, *defined on a Boolean-like ring* B, *be a commutative operation, that is,* $x\circ y=y\circ x$ *for all* $x, y\in B$ *(or* $f(x, y)=f(y, x)$ *for all* $x, y\in B$*), is that* $f(0, 1) = f(1, 0)$, $a_0\equiv a_4$ (mod Q), $a_1\equiv a_6$ (mod Q), $a_2\equiv a_5$ (mod Q) *and* $a_3\equiv a_7$ (mod Q) *where* $f(x, y)$ *is given as in Theorem* 4.

Proof. Assume that $x\circ y=y\circ x$. Then, as a consequence of Theorem 5, we can equate coefficients to obtain the condition.

Assume now that the condition holds. By substituting, for some of the coefficients, other representatives of the same equivalence class modulo Q (for example, substitute a_0 for a_4) it is easily seen that $x\circ y=y\circ x$.

A question arises. Given a Boolean-like ring is it possible to define on this ring binary operations, in terms of the given operations, in order to form a Boolean ring? Since every Boolean ring is commutative and every element is idempotent, the following theorem settles the question.

THEOREM 7. *Let* B *be a Boolean-like ring such that for some nilpotent element* $n\in B$, $n\neq 0$ *(that is,* B *is not Boolean). Then there does not exist a commutative idempotent binary operation* $\circ$ *defined on* B.

Proof. Let $x\in B$ with $x=x_i+x_n$. Then $x\circ x=f(1, 1)x_i+f(0, 0)x_i' +(a_0+a_4)x_ix_n +(a_3+a_7)x_i'x_n$. Assume now that $x\circ x=x$. From Theorem 1, $x=x_i+x_n=1x_i +0x_i' +1x_ix_n+1x_i'x_n$. Hence $f(1, 1)x_i+f(0, 0)x_i' +(a_0+a_4)x_ix_n+(a_3+a_7)x_i'x_n =1x_i+0x_i' +1x_ix_n+1x_i'x_n$. Thus, by Theorem 3, $a_3+a_7\equiv 1$ (mod Q), but, since $\circ$ is a commutative operation, $a_3\equiv a_7$ (mod Q), which implies that $a_3+a_7\equiv 0$ (mod Q). Therefore $1\equiv 0$ (mod Q). Hence $1n=0$ for all nilpotent elements and $1n=n$. Thus $n=0$ for all nilpotent elements which would contradict the assumption that there exists an $n\in B$ such that $n\neq 0$.

THEOREM 8. *Any Boolean-like function of* m *variables is of the form*

$$f(x_1, x_2, \cdots, x_m) = \sum_{(e)} f(e_1, e_2, \cdots, e_m)x_{1_i}^{e_1}x_{2_i}^{e_2}\cdots x_{m_i}^{e_m}$$
$$+\sum_{k=1}^{m}\left\{\sum_{(e)} a_{s_k}x_{1_i}^{e_1}x_{2_i}^{e_2}\cdots x_{m_i}^{e_m}\right\}x_{k_n}$$

where the summation $\sum_{(e)}$ *extends over all combinations* (e) *on* m *zeros and ones and* $x_{p_i}^0=x_{p_i}'$, $x_{p_i}^1=x_{p_i}$. *Also* x_{p_i}, x_{p_n} *are the idempotent and nilpotent parts, respectively, of* x_p *and* $a_{s_k}=a_{(e_1,e_2,\cdots e_m)k}$ *where each* a_{s_k} *is an idempotent element of the Boolean-like ring.*

Proof. The proof follows the outline of Theorem 1. It should be noted in conclusion that if the ring is Boolean then each $x_{p_n}=0$ for all $p=1, 2, \cdots, m$ and the function reduces to the Boolean function as given in [2].

References

1. A. L. Foster, The theory of Boolean-like rings, Trans. Amer. Math. Soc., 59 (1946) 166–187.
2. F. E. Hohn, Applied Boolean Algebra: An Elementary Introduction, Macmillan, New York, 1960.

ON THE NUMBER OF SOLUTIONS OF A CONGRUENCE

K. S. WILLIAMS, University of Toronto

1. Notation. Let p be a prime, $p\equiv 1 \pmod 3$. Then the pair of integers A and B are uniquely determined by the relations:

$$4p = A^2 + 27B^2 \qquad A \equiv 1 \pmod 3, \quad B > 0.$$

We shall write $\theta_1=\frac{1}{2}(A+3B\sqrt{-3})$, $\theta_2=\bar{\theta}_1$ so $\theta_1\theta_2=p$. The complex cube roots of unity will be written w and w^2, where $w=\frac{1}{2}(-1+\sqrt{-3})$. Further the principal character (mod p) will be denoted by χ_0 and the two nonprincipal cubic ones by χ_1 and χ_2. In order to distinguish between these we have:

$$\chi_1(\alpha) = 1, w, w^2 \text{ according as } \alpha^{(p-1)/3} \equiv 1, w, w^2 \pmod{\theta_1}$$

$$\chi_2(\alpha) = 1, w^2, w \text{ according as } \alpha^{(p-1)/3} \equiv 1, w, w^2 \pmod{\theta_1}.$$

Hence we have $\chi_1^2=\chi_2$, $\chi_2^2=\chi_1$ and $\chi_1\chi_2=\chi_0$. Finally we write $e(t)$ for $\exp(2\pi i t p^{-1})$ and define $\tau_i(\alpha)\,(i=1, 2)$ by

$$\tau_i(\alpha) = \sum_{x=1}^{p-1} \chi_i(x)e(\alpha x)$$

with $\tau_i=\tau_i(1)$. So we have the following relations

$$\tau_1\tau_2 = p, \qquad \tau_1^3 = p\theta_1, \qquad \tau_2^3 = p\theta_2 \tag{1.1}$$

and $\tau_1(\alpha)=\chi_2(\alpha)\tau_1$, $\tau_2(\alpha)=\chi_1(\alpha)\tau_2$.

2. Introduction. The above notation is essentially that used by Eckford Cohen in [1]. We shall find in this article an expression for the number of solutions N_n of the congruence:

$$a_1x_1^3 + a_2x_2^3 + \cdots + a_nx_n^3 + b \equiv 0 \pmod p, \quad \text{where} \quad p \nmid \prod_{i=1}^{n} a_i.$$

3. We first need a simple lemma.

LEMMA. *If* $p\nmid a$ *then* $\sum_{x=0}^{p-1} e(ax^3) = \tau_2\chi_1(a)+\tau_1\chi_2(a)$.

Proof.

$$\sum_{x=0}^{p-1} e(ax^3) = 1 + \sum_{y=1}^{p-1} e(ay)N(x: x^3 \equiv y) = 1 + \sum_{y=1}^{p-1} e(ay)[\chi_0(y) + \chi_1(y) + \chi_2(y)]$$

$$= \tau_1(a) + \tau_2(a) = \tau_1\chi_2(a) + \tau_2\chi_1(a) \qquad \text{using (1.1).}$$

4. Now

$$N_n = \frac{1}{p}\sum_{t=0}^{p-1}\sum_{x_1=0}^{p-1} \cdots \sum_{x_n=0}^{p-1} e\{t(a_1x_1^3 + \cdots + a_nx_n^3 + b)\}$$

$$= p^{n-1} + \frac{1}{p}\sum_{t=1}^{p-1} e(bt)\left\{\sum_{x_1=0}^{p-1} e(a_1tx_1^3)\right\} \cdots \left\{\sum_{x_n=0}^{p-1} e(a_ntx_n^3)\right\}$$

$$= p^{n-1} + \frac{1}{p}\sum_{t=1}^{p-1} e(bt)\{\tau_2\chi_1(a_1t) + \tau_1\chi_2(a_1t)\} \cdots \{\tau_2\chi_1(a_nt) + \tau_1\chi_2(a_nt)\}$$

using the lemma.

Now set

$$u_i = \tau_2\chi_1(a_i), \qquad v_i = \tau_1\chi_2(a_i) \qquad \text{for } i = 1, 2, \cdots, n. \tag{4.1}$$

Therefore

$$N_n = p^{n-1} + \frac{1}{p}\sum_{t=1}^{p-1} e(bt)\prod_{i=1}^{n} (u_i\chi_1(t) + v_i\chi_2(t)).$$

Define $x_n \equiv x_n(u_1, u_2, \cdots, u_n; v_1, v_2, \cdots, v_n) \equiv x_n(\underline{u}^{(n)}, \underline{v}^{(n)})$, similarly define y_n, z_n (noting that $x_n(\underline{v}^{(n)}, \underline{u}^{(n)})$ denotes $x_n(\underline{u}^{(n)}, \underline{v}^{(n)})$, where u_i has been replaced by v_i $(i=1, 2, \cdots, n)$ and vice-versa) by

$$x_n\chi_0(t) + y_n\chi_1(t) + z_n\chi_2(t) = \prod_{i=1}^{n} (u_i\chi_1(t) + v_i\chi_2(t)).$$

Thus

$$N_n = p^{n-1} + \frac{1}{p}\sum_{t=1}^{p-1} e(bt)[x_n\chi_0(t) + y_n\chi_1(t) + z_n\chi_2(t)].$$

Therefore,

$$N_n = p^{n-1} + \frac{x_n}{p}(p-1) \qquad \text{if } b \equiv 0 \pmod{p} \tag{4.2}$$

or

$$N_n = p^{n-1} - \frac{x_n}{p} + \frac{y_n}{p}\chi_2(b)\tau_1 + \frac{z_n}{p}\chi_1(b)\tau_2 \qquad \text{if } b \not\equiv 0 \pmod{p}. \tag{4.3}$$

5. We now define a symbol $[m, n-m]$ in terms of which we shall give x_n, y_n, z_n.

DEFINITION. *Let m and n be fixed integers such that* $0 \leqq m \leqq n$. *Define* $[m, n-m]$ *to be* $\sum u \cdots uv \cdots v$, *where there are m u's and* $(n-m)$ *v's and the subscripts in some order form the sequence* $1, 2, \cdots, n$. *In all there will be* $\binom{n}{m}$ *terms.*

EXAMPLES: $[0, 3] = v_1v_2v_3$, $[2, 1] = u_1u_2v_3 + u_1v_2u_3 + v_1u_2u_3$.

We understand $[0, 0]$ to mean 1. We thus have the following identity:

$$u_{n+1}[m, n-m] + v_{n+1}[m+1, n-m-1] = [m+1, n-m]. \tag{5.1}$$

6. Now

$$\begin{aligned} x_n\chi_0(t) &+ y_n\chi_1(t) + z_n\chi_2(t) \\ &= (u_n\chi_1(t) + v_n\chi_2(t))(x_{n-1}\chi_0(t) + y_{n-1}\chi_1(t) + z_{n-1}\chi_2(t)) \\ &= (v_ny_{n-1} + u_nz_{n-1})\chi_0(t) + (u_nx_{n-1} + v_nz_{n-1})\chi_1(t) + (v_nx_{n-1} + u_ny_{n-1})\chi_2(t). \end{aligned}$$

Thus

$$\begin{aligned} x_n(\underline{u}^{(n)}, \underline{v}^{(n)}) &= v_ny_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}) + u_nz_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}) \\ y_n(\underline{u}^{(n)}, \underline{v}^{(n)}) &= u_nx_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}) + v_nz_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}) \\ z_n(\underline{u}^{(n)}, \underline{v}^{(n)}) &= v_nx_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}) + u_ny_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}). \end{aligned}$$

For completeness we define $x_0 = 1$; $y_0 = 0$; $z_0 = 0$. It is straightforward to show by induction that x_n is symmetric; that is,

$$x_n(\underline{u}^{(n)}, \underline{v}^{(n)}) = x_n(\underline{v}^{(n)}, \underline{u}^{(n)}).$$

It then follows immediately that $y_n(u^{(n)}, \underline{v}^{(n)}) = z_n(\underline{v}^{(n)}, \underline{u}^{(n)})$. Now the difference equations simplify and we obtain

$$x_n(\underline{u}^{(n)}, \underline{v}^{(n)}) = v_ny_{n-1}(\underline{u}^{(n-1)}, \underline{v}^{(n-1)}) + u_ny_{n-1}(\underline{v}^{(n-1)}, \underline{u}^{(n-1)}), \tag{6.1}$$

$$\begin{aligned} y_n(\underline{u}^{(n)}, \underline{v}^{(n)}) = {} & u_nv_{n-1}y_{n-2}(\underline{u}^{(n-2)}, \underline{v}^{(n-2)}) + u_nu_{n-1}y_{n-2}(\underline{v}^{(n-2)}, \underline{u}^{(n-2)}) \\ & + v_ny_{n-1}(\underline{v}^{(n-1)}, \underline{u}^{(n-1)}), \end{aligned} \tag{6.2}$$

$$z_n(\underline{u}^{(n)}, \underline{v}^{(n)}) = y_n(\underline{v}^{(n)}, \underline{u}^{(n)}). \tag{6.3}$$

7. Let us define for convenience: $t = [m/6]$,

$$\begin{aligned} f(m) &= 1 \qquad m \equiv 1, 4 \pmod 6 \\ &= 0 \qquad m \equiv 0, 2, 3, 5 \pmod 6 \\ g(m) &= 1 \qquad m \equiv 0, 1, 2 \pmod 6 \\ &= 0 \qquad m \equiv 3, 4, 5 \pmod 6 \\ h(m) &= 0 \qquad m \equiv 0 \pmod 3 \\ &= 2 \qquad m \equiv 1 \pmod 3 \\ &= 1 \qquad m \equiv 2 \pmod 3. \end{aligned}$$

Now we show that the solution of the difference equations is given by

$$(7.1)\quad x_m = \sum_{l=0}^{t-f(m)} [3l + h(m), m - 3l - h(m)] + \sum_{l=0}^{t-g(m)} [m - 3l - h(m), 3l + h(m)],$$

$$(7.2)\quad \begin{aligned} y_m = &\sum_{l=0}^{t-f(m+1)} [3l + h(m-2), m - 3l - h(m-2)] \\ &+ \sum_{l=0}^{t-g(m)} [m - 3l - h(m-1), 3l + h(m-1)], \end{aligned}$$

$$(7.3)\quad \begin{aligned} z_m = &\sum_{l=0}^{t-g(m)} [3l + h(m-1), m - 3l - h(m-1)] \\ &+ \sum_{l=0}^{t-f(m+1)} [m - 3l - h(m-2), 3l + h(m-2)]. \end{aligned}$$

We begin the inductive proof. We assume that the expressions for the y_m are true for $m=0, 1, 2, \cdots, 6n-1$ and deduce that they are valid for $m=6n$, $6n+1, \cdots, 6n+5$. It is easily verified that they are indeed valid for $m=0$, 1, 2, 3, 4, 5. We illustrate the inductive step from $m=6n-1$ to $m=6n$. The rest are similar. From the recurrence relations (6.1) and (6.3) we then have x_n and z_n. Now from (6.2),

$$\begin{aligned} &y_{6n}(\underline{u}^{(6n)}, \underline{v}^{(6n)}) \\ &= u_{6n}v_{6n-1}y_{6n-2}(\underline{u}^{(6n-2)}, \underline{v}^{(6n-2)}) + u_{6n}u_{6n-1}y_{6n-2}(\underline{v}^{(6n-2)}, \underline{u}^{(6n-2)}) \\ &\quad + v_{6n}y_{6n-1}(\underline{v}^{(6n-1)}, \underline{u}^{(6n-1)}) \\ &= u_{6n}v_{6n-1}\sum_{l=0}^{n-1} [3l+1, 6n-3l-3] + u_{6n}v_{6n-1}\sum_{l=0}^{n-1} [6n-3l-2, 3l] \\ &\quad + u_{6n}u_{6n-1}\sum_{l=0}^{n-1} [3l, 6n-3l-2] + u_{6n}u_{6n-1}\sum_{l=0}^{n-1} [6n-3l-3, 3l+1] \\ &\quad + v_{6n}\sum_{l=0}^{n-1} [6n-3l-1, 3l] + v_{6n}\sum_{l=0}^{n-1} [3l+2, 6n-3l-3] \\ &= u_{6n}\sum_{l=0}^{n-1} \{u_{6n-1}[3l, 6n-3l-2] + v_{6n-1}[3l+1, 6n-3l-3]\} \\ &\quad + u_{6n}\sum_{l=0}^{n-1} \{u_{6n-1}[6n-3l-3, 3l+1] + v_{6n-1}[6n-3l-2, 3l]\} \\ &\quad + v_{6n}\sum_{l=0}^{n-1} [6n-3l-1, 3l] + v_{6n}\sum_{l=0}^{n-1} [3l+2, 6n-3l-3] \end{aligned}$$

$$= \sum_{l=0}^{n-1} \{u_{6n}[3l+1,\ 6n-3l-2] + v_{6n}[3l+2,\ 6n-3l-3]\}$$

$$+ \sum_{l=0}^{n-1} \{u_{6n}[6n-3l-2,\ 3l+1] + v_{6n}[6n-3l-1,\ 3l]\} \qquad \text{using (5.1)}$$

$$= \sum_{l=0}^{n-1} [3l+2,\ 6n-3l-2] + \sum_{l=0}^{n-1} [6n-3l-1,\ 3l+1] \quad \text{using (5.1) again.}$$

This is correct since $t=n$, $f(m+1)=1$, $g(m)=1$, $h(m-1)=1$, $h(m-2)=2$. This completes the proof.

8. We now define a new symbol, similar to $[m,\ n-m]$, namely $[\chi_1(a\colon m)\chi_2(a\colon n-m)]$ for positive integers n, m such that $0 \leqq m \leqq n$. The explanation of this symbol is perhaps best illustrated by two examples:

Example (i)

$$[\chi_1(a\colon 1)\chi_2(a\colon 2)] = \chi_1(a_1)\chi_2(a_2a_3) + \chi_1(a_2)\chi_2(a_3a_1) + \chi_1(a_3)\chi_2(a_1a_2).$$

Example (ii)

$$[\chi_1(a\colon 0)\chi_2(a\colon 3)] = \chi_2(a_1a_2a_3).$$

In general there will be $\binom{n}{m}$ terms.

9. Conclusion. If $b \equiv 0 \pmod{p}$, using (1.1), (4.1), (4.2) and (7.1) we find that

$$N_n = p^{n-1} + \frac{(p-1)}{p}\left\{\tau_1^n \sum_{l=0}^{t-f(n)} \left(\frac{\theta_2}{\theta_1}\right)^l \left(\frac{\tau_2}{\tau_1}\right)^{h(n)} [\chi_1(a\colon 3l+h(n))\chi_2(a\colon n-3l-h(n))]\right.$$
$$\left. + \tau_2^n \sum_{l=0}^{t-g(n)} \left(\frac{\theta_1}{\theta_2}\right)^l \left(\frac{\tau_1}{\tau_2}\right)^{h(n)} [\chi_1(a\colon n-3l-h(n))\chi_2(a\colon 3l+h(n))]\right\}.$$

In particular for $n=1, 2, 3$ we have the familiar results:

$$N_1 = 1,$$
$$N_2 = p + (p-1)(\chi_1(a_1)\chi_2(a_2) + \chi_1(a_2)\chi_2(a_1)),$$
$$N_3 = p^2 + (p-1)(\theta_2\chi_1(a_1a_2a_3) + \theta_1\chi_2(a_1a_2a_3)).$$

If $a_i = 1$ $(i = 1, 2, \cdots, n)$ then the expression for N_n simplifies to

$$N_n = p^{n-1} + \frac{(p-1)}{3p}\left[(\tau_1+\tau_2)^n + (\omega\tau_1+\omega^2\tau_2)^n + (\omega^2\tau_1+\omega\tau_2)^n\right].$$

If $b \not\equiv 0 \pmod{p}$, using (1.1), (4.1), (4.3), (7.1), (7.2), (7.3) we have

$$N_n = p^{n-1}$$
$$+ \frac{\chi_1(b)}{\tau_1}\left\{\tau_1^n \sum_{l=0}^{t-g(n)} \left(\frac{\theta_2}{\theta_1}\right)^l \left(\frac{\tau_2}{\tau_1}\right)^{h(n-1)} [\chi_1(a\colon 3l+h(n-1))\chi_2(a\colon n-3l-h(n-1))]\right.$$

$$+ \tau_2^n \sum_{l=0}^{t-f(n+1)} \left(\frac{\theta_1}{\theta_2}\right)^l \left(\frac{\tau_1}{\tau_2}\right)^{h(n-2)} [\chi_1(a:n-3l-h(n-2))\chi_2(a:3l+h(n-2))]\Bigg\}$$

$$+ \frac{\chi_2(b)}{\tau_2}\Bigg\{\tau_1^n \sum_{l=0}^{t-f(n+1)} \left(\frac{\theta_2}{\theta_1}\right)^l \left(\frac{\tau_2}{\tau_1}\right)^{h(n-2)} [\chi_1(a:3l+h(n-2))\chi_2(a:n-3l-h(n-2))]$$

$$+ \tau_2^n \sum_{l=0}^{t-g(n)} \left(\frac{\theta_1}{\theta_2}\right)^l \left(\frac{\tau_1}{\tau_2}\right)^{h(n-1)} [\chi_1(a:n-3l-h(n-1))\chi_2(a:3l+h(n-1))]\Bigg\}$$

$$- \frac{1}{p}\Bigg\{\tau_1^n \sum_{l=0}^{t-f(n)} \left(\frac{\theta_2}{\theta_1}\right)^l \left(\frac{\tau_2}{\tau_1}\right)^{h(n)} [\chi_1(a:3l+h(n))\chi_2(a:n-3l-h(n))]$$

$$+ \tau_2^n \sum_{l=0}^{t-g(n)} \left(\frac{\theta_1}{\theta_2}\right)^l \left(\frac{\tau_1}{\tau_2}\right)^{h(n)} [\chi_1(a:n-3l-h(n))\chi_2(a:3l+h(n))]\Bigg\}.$$

With $n=1, 2, 3$ we have the known results [2]:

$$N_1 = 1 + \chi_2(a_1)\chi_1(b) + \chi_1(a_1)\chi_2(b),$$

$$N_2 = p + \theta_2\chi_1(a_1a_2b) + \theta_1\chi_2(a_1a_2b) - \chi_1(a_1)\chi_2(a_2) - \chi_1(a_2)\chi_2(a_1),$$

$$N_3 = p^2 + p(\chi_1(a_1a_2)\chi_2(a_3b) + \chi_1(a_2a_3)\chi_2(a_1b) + \chi_1(a_3a_1)\chi_2(a_2b) + \chi_1(a_1b)\chi_2(a_2a_3) + \chi_1(a_2b)\chi_2(a_3a_1) + \chi_1(a_3b)\chi_2(a_1a_2)) - (\theta_2\chi_1(a_1a_2a_3) + \theta_1\chi_2(a_1a_2a_3)).$$

If $a_i=1$ $(i=1, 2, \cdots, n)$ the formula becomes:

$$N_n = p^{n-1} + \frac{\chi_1(b)}{3\tau_1}[(\tau_1+\tau_2)^n + \omega^2(\omega\tau_1+\omega^2\tau_2)^n + \omega(\omega^2\tau_1+\omega\tau_2)^n]$$

$$+ \frac{\chi_2(b)}{3\tau_2}[(\tau_1+\tau_2)^n + \omega(\omega\tau_1+\omega^2\tau_2)^n + \omega^2(\omega^2\tau_1+\omega\tau_2)^n]$$

$$- \frac{1}{3p}[(\tau_1+\tau_2)^n + (\omega\tau_1+\omega^2\tau_2)^n + (\omega^2\tau_1+\omega\tau_2)^n].$$

The author wishes to thank Mr. C. F. Dunkl, now at the University of Wisconsin, for a lively and stimulating conversation in connection with this article.

References

1. Eckford Cohen, Representations by cubic congruences. Proc. Nat. Acad. Sci. vol. 39, no. 2, 119–121, Feb. 1953.

2. ———, The number of solutions of certain cubic congruences, Pacific Journal of Mathematics, vol. 5, Suppl. 2, 877–886, 1955.

Editorial Note: It has been brought to the attention of the editors by Dr. M. G. Beumer of the Hague, Netherlands, that the results in the paper by Q. A. M. M. Yahya "On the generalization of Hilbert's inequality," this MONTHLY, 72(1965) 518–520, are all contained in a paper by Fu Cheng Hsiang, "An inequality for finite sequences," Math. Scandinavica, 5(1957) 12–14.

SOME CHARACTERISTIC PROPERTIES OF THE FAREY SERIES

JEAN A. BLAKE, Oakwood College, Huntsville, Alabama

DEFINITION. *The Farey series F_n of order n is the ascending series of irreducible fractions between* 0 *and* 1 *whose denominators do not exceed n. We assume throughout that $n>1$.*

Thus h/k belongs to F_n if $0\leqq h\leqq k\leqq n$, $(h, k)=1$; the numbers 0 and 1 are included in the forms 0/1 and 1/1. For example:

$$F_{10} = 0/1, 1/10, 1/9, 1/8, 1/7, 1/6, 1/5, 2/9, 1/4, 2/7, 3/10, 1/3, 3/8, 2/5, 3/7, 4/9, 1/2, 5/9, 4/7, 3/5, 5/8, 2/3, 7/10, 5/7, 3/4, 7/9, 4/5, 5/6, 6/7, 7/8, 8/9, 9/10, 1/1.$$

Four well-known theorems dealing with the Farey series are (see [1]):

THEOREM A. *If h/k and h'/k' are two successive terms of F_n, then*

$$kh' - hk' = 1.$$

THEOREM B. *If h/k, h''/k'', and h'/k' are three successive terms of F_n, then*

$$h''/k'' = (h + h')/(k + k').$$

THEOREM C. *If h/k and h'/k' are two successive terms of F_n, then*

$$(k + k') > n.$$

THEOREM D. *No two successive terms of F_n have the same denominator.*

Purpose of this paper. The primary purpose of this paper is to confirm a conjecture of Harold L. Aaron [2].

THEOREM I (AARON'S CONJECTURE). *The sum of the numerators of the fractions of a Farey series F_n is equal to one-half the sum of the denominators of these fractions.*

In addition to this, I shall establish the following:

THEOREM II. *In F_n the denominator of the immediate predecessor and immediate successor of* 1/2 *is equal to the greatest odd integer $\leqq n$.*

In order to establish Theorem I, I shall prove a few preliminary lemmas.

LEMMA 1. *The sum of the two fractions between* 0 *and* 1 *belonging to F_n and equidistant from* 1/2 *is equal to* 1 (*such fraction pairs are hereafter termed complementary*).

Proof. Suppose that h/n is a number of this series which is less than 1/2 and such that $(h, n)=1$. Comparing the corresponding number on the other side of

$1/2$ we find $(n-h)/n$. For it to belong to F_n it is necessary that $((n-h), n)=1$.

Suppose, on the contrary, that $(n-h)$ and n are not relatively prime. Then, $n-h=rd$ and $n=rd+h$. Also, $n=sd$ and hence $sd=rd+h$; therefore, d divides h, and consequently d divides $(n-h)$ for d divides n. This contradicts the fact that $(n-h)/n$ was in its lowest terms (definition of terms in a Farey series), and therefore $((n-h), n)=1$. We verify directly that h/n and $(n-h)/n$ are equidistant from $1/2$.

LEMMA 2. *In F_n, the immediate predecessor of* $1/2$ *and its immediate successor are complementary.*

Proof. Let g/m be the immediate predecessor of $1/2$. If there is another term immediately succeeding $1/2$ and belonging to F_n, then by Theorem B we know that if the fractions are g/m, $1/2$, x/y (arranged in order of magnitude), then $1/2=(g+x)/(m+y)$. Now to prove that $g/m+x/y=1$; x/y must be in its lowest terms if it is to belong to F_n. Two fractions whose sum is 1 are in their lowest terms if and only if their denominators are equal, as was brought out in Lemma 1.

Hence, if $g/m+x/y=1$ we must have $y=m$, and consequently $x=(m-g)$. But, g/m was the immediate predecessor of $1/2$, and $x/y=(m-g)/m$ the immediate successor; therefore, the Lemma is proved.

LEMMA 3. *If g/m is the immediate predecessor of r/s, then $(m-g)/m$ is the immediate successor of $(s-r)/s$ and vice versa.*

Proof. By Lemma 2 it was proved that the immediate predecessor of $1/2$ and its immediate successor are complementary. If we take the term r/s immediately preceding g/m and belonging to F_n then by Lemma 1 we can find the complement of r/s, seeing that that is equidistant from $1/2$. Thus we find that the complement to r/s, $(s-r)/s$, has one term between it and $1/2$, i.e. $(m-g)/m$.

By first starting with $(s-r)/s$ we could by the same reasoning find r/s; thus the lemma is established.

Proof of Theorem I. Utilizing Lemmas 1, 2 and 3 we can now consider a typical Farey series F_n which has the form:

$$0/1, \cdots, h/n, g/m, r/s, \cdots, (s-r)/s, (m-g)/m, (n-h)/n, \cdots, 1/1.$$

Here the sum of the numerators divided by the sum of the denominators is:

$$\frac{0+\cdots+h+g+r+\cdots+1+\cdots+s-r+m-g+n-h+\cdots+1}{1+2n+2m+2s+2+\cdots+1},$$

or

$$\frac{0+h-h+g-g+r-r+\cdots+1+n+m+s+\cdots+1}{1+2n+2m+2s+2+\cdots+1},$$

and hence

$$\frac{n + m + s + \cdots + 2}{2n + 2m + 2s + \cdots + 4},$$

thereby proving the theorem.

Proof of Theorem II. Suppose that n is odd; then we have to consider F_{2n+1}. The fraction less than and differing least from $1/2 = (n+1/2)/(2n+1)$ with denominator $(2n+1)$ is $n/(2n+1)$. Suppose that there is a fraction of F_{2n+1} between $n/(2n+1)$ and $1/2$. This fraction has to have a denominator which is less than $(2n+1)$. Suppose it is odd. Then we get:

$$n/(2n + 1) < h/(2h + 1) < 1/2.$$

From this we see easily that $n<h$, which contradicts $(2n+1)>(2h+1)$. Clearing of fractions gives: $(2h+1)n<(2n+1)h$, that is, $2hn+n<2hn+h$.

For an even denominator, we have

$$n/(2n + 1) < (h - 1)/2h < 1/2.$$

Again $(2n+1)>2h$, and $2hn<(2n+1)(h-1)$, $2n+1<h$, which contradicts $(2n+1)>2h$. If, therefore, n is odd, the denominator in F_n of the immediate predecessor of $1/2$ is equal to n.

Now, consider F_{2n+2}. As can be readily observed, to construct F_{2n+2} having given F_{2n+1}, requires the addition of the irreducible fractions of denominator $(2n+2)$ to the series F_{2n+1}. The fraction nearest to $1/2$ with this denominator is $n/(2n+2)$. For F_{2n+1} we discovered that $n/(2n+1)<1/2$. The question now is: does $n/(2n+2)$ lie between $n/(2n+1)$ and $1/2$?

Consider $n/(2n+1)<n/(2n+2)<1/2$. This is obviously impossible because of the relative size of the denominators. Hence $n/(2n+1)$ is closest to $1/2$ in F_{2n+2}. We may, therefore, conclude that the immediate predecessor of $1/2$ is the greatest odd integer $\leqq n$.

A similar reasoning establishes the nearest term immediately succeeding $1/2$ in F_n is the greatest odd integer $\leqq n$.

References

1. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, London, 1954, pp. 23–24.

2. Harold L. Aaron, Observation made in Number Theory class during school year 1962–1963, Howard University, Washington, D.C.

A PROOF OF 4-COLORING THE EDGES OF A CUBIC GRAPH

E. L. Johnson, Operations Research Center, University of California, Berkeley

The definitions are those of [3]. A cubic graph is a finite, undirected, regular three-degree graph without loops, but with multiple edges permitted. A graph is called strongly connected if it is connected and not 1-edge connected.

Coloring the edges of a cubic graph is of special interest because the 4-color conjecture is equivalent to 3-coloring the edges of a planar, strongly connected cubic graph [4]. The purpose of this paper is to present a simple proof that the edges of any cubic graph can be 4-colored. This result follows from a theorem on coloring the vertices [1] and is a special case of Shannon's theorem [5]. The result in a strongly connected cubic graph follows from Peterson's theorem on existence of a perfect matching in a strongly connected cubic graph. Frink's method of proving Peterson's theorem in [2] can be used, however, to 4-color the edges so that no color is changed once it is assigned. In contrast, both Shannon's proof and Frink's proof of Peterson's theorem require changes along alternating paths.

THEOREM. *The edges of a cubic graph can be 4-colored.*

Proof. If the theorem is false, then there is a smallest cubic graph G for which it fails. Clearly, G is connected and contains no 3-edges.

Suppose that G is strongly connected. If G contains a 2-edge,

FIG. 1

then form G' from G by eliminating vertices v_2 and v_3 and the edges incident to them, and adjoining edge $[v_1, v_4]$. Because G is strongly connected $v_1 \neq v_4$ so G' is a smaller cubic graph than G, and hence its edges can be 4-colored. But then the edges of G can be 4-colored by coloring $[v_1, v_2]$ and $[v_3, v_4]$ the same color as $[v_1, v_4]$ in G', and coloring the two edges $[v_2, v_3]$ any two of the remaining three colors. Hence, if G is strongly connected, then it contains no 2-edge.

Still supposing G to be strongly connected and now containing no 2-edge, if G contains a triangle,

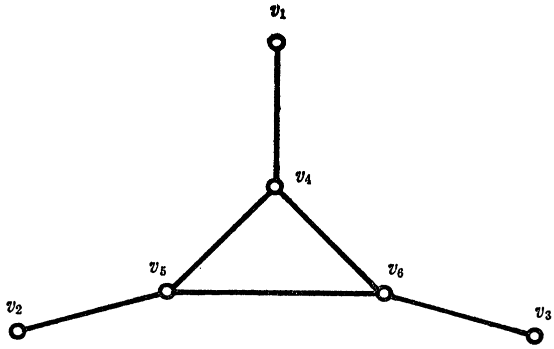


FIG. 2

then form G' by eliminating v_4, v_5, and v_6 and the edges incident to them, and adjoining vertex v' and edges $[v_1, v']$, $[v_2, v']$, and $[v_3, v']$. Then G' is a cubic graph so its edges can be 4-colored. The edges of G can also be 4-colored by coloring $[v_1, v_4]$ and $[v_5, v_6]$ the same as $[v_1, v']$; $[v_2, v_5]$ and $[v_4, v_6]$ the same as $[v_2, v']$; and $[v_3, v_6]$ and $[v_4, v_5]$ the same as $[v_3, v']$. Hence, G contains no triangles.

Hence, G is either strongly connected with neither 2-edges nor triangles, or G is 1-edge connected. In the first case, let $[v_1, v_2]$ be any edge. In the second case, there is an edge $[v_1, v_2]$ such that its removal causes G to break into two connected components. Consider the two other incident edges at v_1 and v_2.

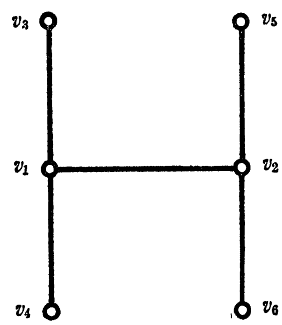


FIG. 3

Form G' by eliminating v_1 and v_2 and all incident edges from G, and adjoining $[v_3, v_5]$ and $[v_4, v_6]$. In either case $v_3 \neq v_5$ and $v_4 \neq v_6$, so no loops are introduced, although multiple edges might be introduced. Therefore, G' is a cubic graph and its edges can be 4-colored. But then the edges of G can also be 4-colored as follows.

If $[v_3, v_5]$ and $[v_4, v_6]$ are colored different colors in G', then color $[v_1, v_3]$ and $[v_2, v_5]$ the same in G as $[v_3, v_5]$ was in G', color $[v_1, v_4]$ and $[v_2, v_6]$ the same in G as $[v_4, v_6]$ was in G', and color $[v_1, v_2]$ either of the two remaining colors.

If $[v_3, v_5]$ and $[v_4, v_6]$ are colored the same colors in G', then color $[v_1, v_3]$ and $[v_2, v_5]$ the same in G as $[v_3, v_5]$ was in G', color $[v_1, v_4]$ the color not incident to v_4 in G', color $[v_2, v_6]$ the color not incident to v_6 in G', and color $[v_1, v_2]$ the remaining color.

REMARK. All of the reductions except that in the very last paragraph permit 3-coloring the edges. The significant problem is to find conditions under which the fourth color is not needed.

References

1. R. L. Brooks, On colouring the nodes of a network, Cambridge Philos. Soc., 37 (1941) 194–197.
2. O. Frink, A proof of Peterson's theorem, Ann. Math., 27 (1926) 491–493.
3. O. Ore, Theory of graphs, Amer. Math. Soc., 1962.

4. O. G. Ringel, Farbungsprobleme auf Flächen und Graphen, Deutschen Verlag der Wissenschaften, Berlin, 1959.

5. C. Shannon, A theorem on coloring the lines of a network, J. Math. Phys., 28 (1949) 148–151.

DIRECT PRODUCTS OF CYCLIC GROUPS

MICHAEL ROSEN, Brown University

Let L be a Galois extension of K whose Galois group is the direct product of n copies of a cyclic group of order m. L is clearly the compositum of n cyclic extensions of K of degree m. We would like to calculate the number of subfields of L which are cyclic extensions of K of degree m. By Galois Theory this is equivalent to calculating the number of subgroups H of G such that G/H is cyclic of order m. We shall answer the more general question of determining the number of subgroups H of G such that G/H is the direct product of $k \leqq n$ copies of a cyclic group of order m.

For a given integer m, let $Z(m)$ be the ring of integers modulo m. Let F be the free $Z(m)$ module of rank n. Then F is isomorphic to G as an abelian group, and there is a one-to-one correspondence between the subgroups of G and the submodules of F. We may thus reformulate our problem as follows: How many submodules $E \subset F$ are there such that F/E is free of rank k? The following result is standard.

LEMMA 1. *If F/E is free then E is a direct summand of F.*

Let C be the set of submodules of F which are direct summands having free complements of rank k. By Lemma 1 our question becomes that of determining the number of elements in C.

LEMMA 2. *A direct summand of F having a free complement is free.*

Proof. Let $F = E \oplus F_1$, where F_1 is free, and let n and n_1 be the ranks of F and F_1 respectively. We will establish the result by applying the fundamental theorem of abelian groups, namely that every finite abelian group is the direct sum of cyclic groups of prime power order, and the number of cyclic summands of each order is unique.

Let $m = p_1^{g_1} \cdots p_r^{g_r}$ be the prime decomposition of m. We have

$$Z(m) = Z(p_1^{g_1}) \oplus \cdots \oplus Z(p_r^{g_r})$$

by the Chinese remainder theorem. Each $Z(p_i^{g_i})$ is cyclic and appears n times in F and n_1 times in F_1. Consequently it appears $n - n_1$ times in E. It follows that E is the direct sum of $n - n_1$ copies of $Z(m)$.

Let $x_1, \cdots, x_n$ be a basis for F and let E_0 be the module generated by $x_1, \cdots, x_{n-k}$. $E_0 \in C$. Let $E \in C$. By Lemma 2, E is free. Thus we can choose a basis $y_1, \cdots, y_n$ of F such that $y_1, \cdots, y_{n-k}$ generate E. Define $\alpha(x_i) = y_i$ and extend α by linearity to an automorphism of F. We have $\alpha(E_0) = E$. Conversely if α is an automorphism then $\alpha(E_0) \in C$.

THEOREM 1. *Let $A(F)$ be the group of automorphisms of F, and let I be the subgroup defined by*

$$I = \{\alpha \in A(F) \mid \alpha(E_0) = E_0\}.$$

The number of elements in C is equal to the index of I in $A(F)$.

Proof. By the preceding remarks the map $\alpha \rightarrow \alpha(E_0)$ maps $A(F)$ onto C. $\alpha, \beta \in A(F)$ have the same image, i.e. $\alpha(E_0) = \beta(E_0)$, if and only if $\beta^{-1}\alpha(E_0) = E_0$, that is if and only if $\beta^{-1}\alpha \in I$. The result follows.

Using the basis $x_1, \cdots, x_n$ we can associate to each element of $A(F)$ a matrix. For $\alpha \in A(F)$ set

$$\alpha(x_i) = \sum_{j=1}^{n} a_{ji}x_j$$

and associate (a_{ji}) to α. The map $\alpha \rightarrow (a_{ji})$ determines an isomorphism between $A(F)$ and the group $Gl_n(Z(m))$ of all n by n matrices with coefficients in $Z(m)$ having for their determinant a unit. Under this isomorphism I goes into the subgroup of $Gl_n(Z(m))$ defined by the property that the rectangle formed by the intersection of the first $n-k$ columns and the last k rows consists entirely of zeros. From the form of these matrices it follows that the order of I is given by

$$m^{k(n-k)} \text{ ord } Gl_{n-k}(Z(m)) \text{ ord } Gl_k(Z(m)).$$

We have reduced our problem to the computation of the order of $Gl_n(Z(m))$. This has already been accomplished by Jordan (see pp. 95–97 in [2]). We shall rederive his result by somewhat different methods. The idea is to reduce the computation to the case in which m is a prime, the answer in this case being well known.

LEMMA 3. *Suppose l divides m. Then there is a natural homomorphism from $Gl_n(Z(m))$ onto $Gl_n(Z(l))$.*

Proof. Let π denote the natural homomorphism from $Z(m)$ onto $Z(l)$. For $(a_{ij}) \in Gl_n(Z(m))$ define $\phi(a_{ij}) = (\pi(a_{ij}))$. To show $(\pi(a_{ij})) \in Gl_n(Z(m))$ notice that det $(\pi(a_{ij})) = \pi$ det (a_{ij}). Since π maps units to units it follows that $(\pi(a_{ij})) \in Gl_n(Z(l))$.

ϕ is easily seen to be a homomorphism. It remains to show that ϕ is onto.

Let $(b_{ij}) \in Gl_n(Z(l))$ and let $(\hat{b}_{ij})$ be an $n \times n$ matrix of integers which reduces to (b_{ij}) modulo l. Let l' be the product of those primes which divide m but not l. By the Chinese remainder theorem there exist integers $\hat{a}_{ij}$ such that $\hat{a}_{ij} \equiv \hat{b}_{ij}(l)$ and $\hat{a}_{ij} \equiv \delta_{ij}(l')$, where $\delta_{ij} = 1$ if $i = j$ and zero otherwise. det $(\hat{a}_{ij})$ is relatively prime to l and l' and consequently to m. Let (a_{ij}) be the reduction of $(\hat{a}_{ij})$ modulo m. Then $(a_{ij}) \in Gl_n(Z(m))$ and $\phi(a_{ij}) = (b_{ij})$. Thus ϕ is onto.

LEMMA 4. *Let $m = p_1^{g_1} \cdots p_r^{g_r}$ be the prime decomposition of m. Then $Gl_n(Z(m))$ is isomorphic to the direct product of the groups $Gl_n(Z(p_i^{g_i}))$.*

Proof. By Lemma 3 there is a homomorphism ϕ_i from $Gl_n(Z(m))$ onto $Gl_n(Z(p_i^{g_i}))$. For $A \in Gl_n(Z(m))$ define

$$\phi(A) = (\phi_1(A), \cdots, \phi_r(A)) \in Gl_n(Z(p_1^{g_1})) \times \cdots \times Gl_n(Z(p_r^{g_r})).$$

ϕ is the required isomorphism. We omit the more or less straightforward details.

LEMMA 5. *Let p be a prime. The order of $Gl_n(Z(p^g))$ is $p^{gn^2}(1-p^{-1}) \cdots (1-p^{-n})$.*

Proof. In Lemma 3 we defined a natural homomorphism from $Gl_n(Z(p^g))$ onto $Gl_n(Z(p))$. The kernel of this homomorphism consists of all matrices of the form $I+M$ where I is the identity of $Gl_n(Z(p^g))$ and M is an arbitrary n by n matrix with elements in $pZ(p^g)$. Since $pZ(p^g)$ has p^{g-1} elements, the kernel has $p^{(g-1)n^2}$ elements. The order of $Gl_n(Z(p))$ is

$$(p^n - 1) \cdots (p^n - p^{n-1}) = p^{n^2}(1 - p^{-1}) \cdots (1 - p^{-n})$$

(see, for example, [1], p. 169). Putting these facts together we see that the order of $Gl_n(Z(p^g))$ is

$$p^{(g-1)n^2} p^{n^2}(1 - p^{-1}) \cdots (1 - p^{-n}) = p^{gn^2}(1 - p^{-1}) \cdots (1 - p^{-n}).$$

THEOREM 2. *The order of $Gl_n(Z(m))$ is*

$$m^{n^2} \prod_{p|m} \prod_{i=1}^{n} (1 - p^{-i}).$$

Proof. This follows immediately from Lemmas 4 and 5.

THEOREM 3. *Let G be the direct product of n copies of a cyclic group of order m. Then the number of sub-groups $H \subset G$, such that G/H is the direct product of $k \leq n$ copies of a cyclic group of order m, is equal to*

$$m^{k(n-k)} \prod_{p|m} \prod_{i=1}^{n-k} \left[\frac{1 - p^{-(i+k)}}{1 - p^{-i}} \right].$$

Proof. By Theorem 1 and the remarks following it, the number we are looking for is the order of $Gl_n(Z(m))$ divided by

$$m^{k(n-k)} \operatorname{ord} Gl_k\,(Z(m)) \operatorname{ord} Gl_{n-k}(Z(m)).$$

The result follows by applying Theorem 2 and simplifying.

This work has been partially supported by the NSF Grant GP-1569.

References

1. E. Artin, Geometric Algebra, Interscience Publ., New York, 1957.
2. C. Jordan, Traité des Substitutions, Gauthier-Villars, Paris, 1870.

Proof. Since the functional M is multiplicative, we have

$$L(f)L(g) \leqq \binom{2m}{m}^{1/2}\binom{2n}{n}^{1/2} M(f)M(g) = \left\{\binom{2m}{m}\binom{2n}{n}\right\}^{1/2} M(fg)$$

$$\leqq \left\{\binom{2m}{m}\binom{2n}{n}\right\}^{1/2} L(fg)$$

by Theorems 1 and 2.

It has also been observed by Mahler [2] that

$$M(f+g) \leqq 2^m\{M(f) + M(g)\} \tag{2}$$

for polynomials $f(z)$ and $g(z)$ of degree at most m and that such inequalities have applications in the theory of diophantine approximations. Since

$$\binom{2m}{m}^{1/2} = \left(\sum_{k=0}^{m}\binom{m}{k}^2\right)^{1/2} < \sum_{k=0}^{m}\binom{m}{k} = 2^m$$

for $m \geqq 1$, the following result is an improvement over (2).

THEOREM 4. *If $f(z)$ and $g(z)$ are polynomials of degree at most m, then*

$$M(f+g) \leqq \binom{2m}{m}^{1/2}\{M(f) + M(g)\}.$$

Proof. Since the sequence $\{\binom{2m}{m}^{1/2}\}$ is monotonic increasing, we have $M(f+g) \leqq L(f+g) \leqq L(f) + L(g) \leqq \binom{2m}{m}^{1/2}\{M(f)+M(g)\}$ by Minkowski's inequality [3, Sec. 12.43] and Theorems 1 and 2.

By using Stirling's formula [3, Sec. 1.87] and calculating a few simple limits we see easily that $\binom{2m}{m}^{1/2} \sim (\pi m)^{-1/4}2^m$. Thus the constants in the above theorem have lower order of magnitude than those in (2).

References

1. K. Mahler, An application of Jensen's formula to polynomials, Mathematica, 7 (1960) 98–100.

2. ———, On the zeros of the derivative of a polynomial, Proc. Roy. Soc. A, 264 (1961) 145–154.

3. E. C. Titchmarsh, Theory of Functions, Oxford, 1939.

4. G. H. Hardy, J. E. Littlewood and G. Polya, Inequalities, Cambridge, 1952.

5. N. J. Fine, Generating functions, 28th Yearbook NCTM, 1963.

THE STRUCTURE OF NEAR RINGS ON A GROUP OF PRIME ORDER

R. A. JACOBSON, Houghton College

In a recent note [1], Clay determined a lower bound for the number of left near rings for the triple $\langle G, +, *\rangle$. An additional theorem enables us to analyze the entire structure of such near rings for all groups where $\langle G, +\rangle$ is of prime order. A natural consequence of such a discussion is an exact expression for the

number of left near rings which verify the results computed in [1] for $n=3, 5, 7$, and also discloses the previous error for $n=2$.

In the following, we let $P=\{0, 1, \cdots, p-1\}$ where p is prime and q is a specific primitive of p. Let N be the multiplicative subgroup of $P\smallsetminus\{0\}$ generated by q^n and rN denote the coset containing the element r. Borrowing from [1], we let $r * 1=b$ define the function $\pi(r)=b$ and recall Theorem II: $*$ is an associative left distributive binary operation if and only if

$$\pi(r)\cdot\pi(s) = \pi(r\cdot\pi(s)), \tag{1}$$

where $\cdot$ implies multiplication, (mod p).

Assuming that $\pi(y)\neq 0$ for some $y\in P$, we have

THEOREM 1. *Let* $n=$g.l.b. $\{x\mid \exists y$ *and a positive integer* x *such that* $y\in P$ *and* $\pi(y)=q^x\}$ *and* $\pi(0)=0$. *If* $*$ *is an associative left distributive binary operation it follows that*:

(i) *if* $\pi(r)\neq 0$, *then* $\pi(r)=q^x$, *where* $n\mid x$;
(ii) *if* $\pi(r)=0$, *then* $\pi(s)=0$, $\forall s\in rN$;
(iii) *if* $\pi(r)\neq 0$, *then* $\forall b\in N$, $\exists s\in rN$ *such that* $\pi(s)=b$.

Proof. In the following, let v be such that $\pi(v)=q^n$.

(i) Let $\pi(r)=q^\alpha$ and $\pi(s)=q^\beta$. From (1), we have $q^{\alpha+\beta}=\pi(r)\cdot\pi(s)=\pi(rq^\beta)$ and the set $X=\{x\mid \exists y$ such that $y\in P$ and $\pi(y)=q^x\}$ is closed under addition, (mod p). Thus, the elements of X are all multiples of the least positive element n.

(ii) If $\pi(r)=0$, we use (1) to write $0=\pi(r)\cdot\pi(v)=\pi(r\cdot\pi(v))=\pi(rq^n)$. A reapplication of (1) gives

$$0 = \pi(rq^n)\cdot\pi(v) = \pi(rq^n\cdot\pi(v)) = \pi(rq^{2n}).$$

It is evident that a continuation of this procedure results in $\pi(rq^{\alpha n})=0$ for all elements $rq^{\alpha n}\in rN$.

(iii) Let $\pi(r)=q^{\alpha n}$. Then employing (1) we have

$$q^{2\alpha n} = \pi(r)\cdot\pi(r) = \pi(r\cdot\pi(r)) = \pi(rq^{\alpha n}).$$

Reapplications of (1) yield

$$q^{3\alpha n} = \pi(rq^{\alpha n})\cdot\pi(r) = \pi(rq^{\alpha n}\cdot\pi(r)) = \pi(rq^{2\alpha n}),$$

and so on, until we have $1=\pi(rq^{\beta n})$ for some element $rq^{\beta n}=t\in rN$. Since $\pi(t)=1$, it follows that $q^n=\pi(t)\cdot\pi(v)=\pi(t\cdot\pi(v))=\pi(tq^n)$ where $tq^n=u\in rN$. Then considering products $q^{2n}=\pi(u)\cdot\pi(u)=\pi(u\pi(u))=\pi(uq^n)$, etc.; we find by induction that

$$\pi(uq^{xn}) = q^{(x+1)n} \tag{2}$$

and the theorem is proved.

REMARK. The structure of $\langle G, *\rangle$ can now be established by noting that whenever the hypothesis of Theorem 1 is satisfied, then any combination of

cosets can have an element u such that $\pi(u)=q^n$; furthermore, the element u in a particular coset can be chosen at random.

Proof. In order to establish the remark we must show that (1) holds for all $r, s\in P$. Since $\pi(0)=0$, it is evident that (1) follows if either $r=0$ or $s=0$. On the other hand, if $rs\neq 0$, then from Theorem 1, there is an element $a\in rN$ such that $\pi(a)\in\{0, q^n\}$. Similarly, there exists an element $b\in sN$ with the same property. Upon letting $r=aq^{xn}$ and $s=bq^{yn}$, we find that (1) becomes

$$\pi(aq^{xn})\cdot\pi(bq^{yn}) = \pi(aq^{xn}\cdot\pi(bq^{yn})). \tag{3}$$

If $\pi(b)=0$, we recall that $\pi(0)=0$ and employ (ii) to establish that

$$\pi(aq^{xn})\cdot\pi(bq^{yn}) = \pi(aq^{xn})\cdot 0 = 0,$$
$$\pi(aq^{xn}\cdot\pi(bq^{yn})) = \pi(aq^{xn}\cdot 0) = \pi(0) = 0.$$

If $\pi(b)\neq 0$, $\pi(a)=0$, we again use (ii) together with (2) and find that

$$\pi(aq^{xn})\cdot\pi(bq^{yn}) = 0\cdot q^{(y+1)n} = 0,$$
$$\pi(aq^{xn}\cdot\pi(bq^{yn})) = \pi(aq^{(x+y+1)n}) = 0.$$

Finally, if $\pi(a)=\pi(b)=q^n$, we use (2) to simplify (3), obtaining on the left side

$$\pi(aq^{xn})\cdot\pi(bq^{yn}) = q^{(x+y+2)n},$$

and on the right side

$$\pi(aq^{xn}\cdot\pi(bq^{yn})) = \pi(aq^{(x+y+1)n}) = q^{(x+y+2)n},$$

and thus (1) holds for all r, s. Since the proof does not depend on the choice of a or b, the remark is verified.

In order to find the total number of possible left near rings, we proceed as follows:

(a) enumerate the possible combinations of cosets of N, that is, $C(n, j)$, $j=0, 1, \cdots, n$;

(b) for each coset the element such that $\pi(u)=q^n$ can be chosen in $(p-1)/n$ ways;

(c) thus, the elements u_j such that $\pi(u_j)=q^n$ can be chosen in $((p-1)/n)^j$ ways for any particular combination $C(n, j)$;

(d) finally, including the two cases $\pi(r)=0$, $\forall r\in P$ and $\pi(r)=1$, $\forall r\in P$ (see Theorem IV, [1]); we find the exact number of left near rings is given by

$$2+\sum_k\left\{\sum_{j=1}^{k} C(k,j)\left(\frac{p-1}{k}\right)^j\right\}; \qquad \text{where } k\mid p-1. \tag{4}$$

The correct value of 3 for $p=2$ follows from (4).

Reference

1. James R. Clay, The near rings on a finite cyclic group, this MONTHLY, 71 (1964) 47–50.

ISOMORPHISMS BETWEEN CERTAIN SUBSEMIGROUPS OF THE POSITIVE INTEGERS UNDER MULTIPLICATION

PAGE PAINTER, University of California, Davis

1. Introduction. It has been proved that there is no algorithm for determining when two semigroups are isomorphic [1], and no algorithm is known for the case in which both semigroups are imbedded in the multiplicative integers. We will discuss isomorphisms among the semigroups

$$M_j = \{z \mid z \geqq j\} \quad \text{and} \quad I_j = \{z \mid z = nj \text{ for some } n \in Z^+\}.$$

We note that both of these subsemigroups are ideals in the multiplicative positive integers. Throughout this discussion, all lower case letters will stand for elements of the set of positive integers Z^+, the binary operation in each semigroup will be ordinary multiplication, and the phrase "$\phi: A \to B$ is an isomorphism" will mean "ϕ is an isomorphism of A onto B."

Now any positive integers j and k define the subsemigroups I_j, I_k, M_j, and M_k. For any two of the above subsemigroups, we will find necessary and sufficient conditions for the existence of an isomorphism between them.

2. Isomorphisms between M_j and M_k. We will prove

THEOREM 1. *$M_j \cong M_k$ if and only if $j=k$.*

Proof. Assume that $\phi: M_j \to M_k$ is an isomorphism. We may assume that $j>k$. If $k=1$, M_k contains the identity, and consequently M_j also contains 1. This gives the contradiction $j=k=1$, and we can therefore restrict our attention to the case $j>k>1$. Define $N_j=\{z \mid 1<z<j\}$ and $N_k=\{z \mid 1<z<k\}$. Note that N_j is not empty. Now let α be an integer such that $2^\alpha \geqq j$. Then for all $z \in N_j$ and for all $z \in N_k$, we have $z^\alpha \in M_j$ and $z^\alpha \in M_k$. Now for any $z \in N_j$, $jz \in M_j$, and therefore $\phi(jz) \in M_k$. Since $z^\alpha \in M_j$, we have $\{\phi(jz)\}^\alpha = \{\phi(j)\}^\alpha \phi(z^\alpha)$. But from this equation we see that $\phi(z^\alpha)$ is a positive integer with a rational α'th root and that consequently $\phi(z^\alpha)=x^\alpha$ for some positive integer x. We will show that $x \in N_k$. Since $1 \notin M_k$, $\phi(z^\alpha) \neq 1$. Now, if $x \geqq k$, then $\phi^{-1}(x)=y \in M_j$. But then we have $\phi(y)=x$, and $\phi(y^\alpha)=x^\alpha=\phi(z^\alpha)$. Hence $y=z$, contradicting $z \in N_j$.

Therefore, we can define $\psi: N_j \to N_k$ by $\psi(z)=x$, where $\phi(z^\alpha)=x^\alpha$. We see that ψ is $1-1$, for if $\psi(y)=\psi(z)$, then $\phi(y^\alpha)=\phi(z^\alpha)$, and we have $y=z$. Since N_j properly contains N_k, however, such a function cannot exist. We conclude that $j=k$.

3. Isomorphisms between I_j and I_k. We will prove

THEOREM 2. *$I_j \cong I_k$ if and only if there exist distinct primes $p_1, p_2, \cdots, p_n$, distinct primes $q_1, q_2, \cdots, q_n$, and integers $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that $j=p_1^{\alpha_1}p_2^{\alpha_2} \cdots p_n^{\alpha_n}$ and $k=q_1^{\alpha_1}q_2^{\alpha_2} \cdots q_n^{\alpha_n}$.*

Proof. Assume that $p_1, p_2, \cdots, p_n$ are distinct primes, that $q_1, q_2, \cdots, q_n$ are distinct primes, and that

$$j = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n} \quad \text{and} \quad k = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_n^{\alpha_n}.$$

We will construct an isomorphism $\phi : I_j \to I_k$.

Define $\phi(p_i) = q_i$. Now let χ be a bijection of the set P of all primes other than $p_1, p_2, \cdots, p_n$ onto the set Q of all primes other than $q_1, q_2, \cdots, q_n$, and extend ϕ by χ. Define $\phi(1) = 1$, and for any integer b, define

$$\phi(b) = \phi(b_1)\phi(b_2) \cdots \phi(b_s),$$

where $b_1 b_2 \cdots b_s$ is the factorization of b into primes. Now ϕ is an isomorphism of the positive integers onto the positive integers, and clearly ϕ restricted to I_j is an isomorphism onto I_k.

For the converse, assume that $\phi : I_j \to I_k$ is an isomorphism. We note that any isomorphism preserves the number of distinct factorizations of an element into the product of a given number of factors. (If two factorizations differ only in the order of the factors, we consider them equal.) Now $\phi(j) = kx$ for some integer x, and $\phi(j^2) = (kx)^2$, but j^2 has only one factorization into the product of two elements of I_j, $j^2 = (j)(j)$. Therefore, $(kx)(kx)$ and $(kx^2)(k)$ cannot be distinct. Consequently, $x = 1$ and $\phi(j) = k$.

Now we wish to use ϕ to construct a mapping of the primes into the primes. To do so, we note that if p is a prime, then $\phi(pj) = qk$ for some integer q. We will show that q is prime. Since pj^2 has only one factorization in I_j, $(pj)(j)$, $\phi(pj^2) = qk^2$ can have only the one factorization $(qk)(k)$. But if $q = q'q''$ and $q' > 1$, $q'' > 1$, then $(q'k)(q''k)$ is a distinct factorization in I_k. Therefore, q must be prime. Now define $\psi(p) = q$, where $\phi(pj) = qk$. We note that ψ is $1-1$, for if $\psi(p) = \psi(p')$, then $\phi(pj) = \phi(p'j)$, and hence $p = p'$.

Now factor j into primes, $j = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}$. We have

$$\begin{aligned}\phi\{(jp_1)^{\alpha_1}(jp_2)^{\alpha_2} \cdots (jp_n)^{\alpha_n}\} &= (\phi(jp_1))^{\alpha_1}(\phi(jp_2))^{\alpha_2} \cdots (\phi(jp_n))^{\alpha_n} \\ &= (kq_1)^{\alpha_1}(kq_2)^{\alpha_2} \cdots (kq_n)^{\alpha_n},\end{aligned}$$

where $q_i = \psi(p_i)$. We also have

$$\begin{aligned}\phi\{(jp_1)^{\alpha_1}(jp_2)^{\alpha_2} \cdots (jp_n)^{\alpha_n}\} &= \phi(j^{\alpha_1+\alpha_2+\cdots+\alpha_n} p^{\alpha_1} p^{\alpha_2} \cdots p^{\alpha_n}) \\ &= \phi(j^{\alpha_1+\alpha_2+\cdots+\alpha_n+1}) = k^{\alpha_1+\alpha_2+\cdots+\alpha_n+1}.\end{aligned}$$

From these equations we see that $k = q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_n^{\alpha_n}$, where the q_i are distinct primes, and this completes the proof.

4. Isomorphisms between M_j and I_k.

COROLLARY. *$M_j \cong I_k$ if and only if $j = k = 1$.*

Proof. Assume that $M_j \cong I_k$ and that $k \geqq 2$. We note that the square of every element other than k in I_k has the distinct factorizations $(kx^2)(k)$ and $(kx)(kx)$, while the square of any prime in M_j has only one factorization $(p)(p)$. Therefore, every prime in M_j maps onto k, which contradicts $M_j \cong I_k$.

If $k=1$, then we have $M_j \cong I_1 = M_1$. But by Theorem 1, M_1 is isomorphic only to M_1. Hence we have $j=k=1$.

Reference

1. Michael O. Rabin, Recursive unsolvability of group-theoretic problems, Ann. of Math. Ser. 2, 67 (1958) 188.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

This department welcomes brief expository articles on topics closely related to classroom experience in courses that are normally available to undergraduate students, from the freshman year through early graduate work. Items of interest to teachers, such as pedagogical tactics, course improvement, new proofs and counterexamples, and fresh viewpoints in general, are invited. All material should be sent to Gertrude Ehrlich, Mathematics Department, University of Maryland, College Park, Maryland 20740.

A GENERAL FORM OF THE REMAINDER IN TAYLOR'S THEOREM

P. R. BEESACK, Carleton University, Ottawa, Canada

In a typical derivation (cf. [1] p. 137) of Taylor's theorem with remainder,

$$(1)\quad f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + R_n(x, a),$$

one usually assumes a remainder of a specified form, such as

$$(2)\qquad R_n(x, a) = K(x, a)(x-a)^n, \quad \text{or} \quad R_n(x, a) = K(x, a)(x-a),$$

leading to the Lagrange or Cauchy form of the remainder, respectively. We shall follow the usual method of proof here, first regarding (1) as an *identity* in the variables (x, a), noting what this implies for the function R_n, and then using these observations to obtain some specific—but rather general—forms of the remainder.

As usual, we assume that $f^{(n-1)}$ is continuous on $I=[a, b]$, and that $f^{(n)}$ exists on $I^0=(a, b)$. Defining the function R_n on $I \times I$ by the equation

$$f(x) = f(t) + \frac{f'(t)}{1!}(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} + R_n(x, t),$$

we see that $\partial^{n-1}R_n/\partial x^{n-1}$ is continuous on $I \times I$, $\partial^n R_n/\partial x^n$ exists on $I^0 \times I$, and $\partial R_n/\partial t$ exists on $I \times I^0$. Moreover, $R_n(x, x) \equiv 0$, $x \in I$. Now, for each $x \in I$, the function F_x, defined for $t \in I$

$$F_x(t) = f(x) - \left\{ f(t) + \frac{f'(t)}{1!}(x-t) + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} + R_n(x, t) \right\},$$

is identically zero. Hence also,

$$F_x'(t) = -\frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} - \frac{\partial R_n(x,t)}{\partial t} \equiv 0. \tag{3}$$

Using the fact that $R_n(x, x) \equiv 0$, we thus obtain the explicit evaluation

$$R_n(x,t) = \int_x^t \frac{\partial R_n(x,s)}{\partial s}\,ds = \int_t^x \frac{f^{(n)}(s)}{(n-1)!}(x-s)^{n-1}\,ds.$$

Hence, in (1), we obtain the exact or integral form of the remainder,

$$R_n(x,a) = \int_a^x \frac{f^{(n)}(s)}{(n-1)!}(x-s)^{n-1}\,ds. \tag{4}$$

(This assumes, of course, that $f^{(n)}$ is integrable on I, as does the usual proof.) As is well known, the remainders (2) may be derived from (4) using the integral mean value theorems.

Returning now to equation (3), it is clear that if R_n is known to be "of a certain form," then R_n can be explicitly determined from (3) without any integration. For example, if

$$R_n(x,t) = K(x)g(x,t), \tag{5}$$

where g is a given function, we would have

$$\frac{\partial R_n}{\partial t} = K(x)g_t(x,t) = -\frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1}.$$

Hence, provided $g_t \neq 0$ on $I \times I$, we could solve this equation for K, and thus determine R_n explicitly in the form

$$R_n(x,t) = -\frac{f^{(n)}(t)}{(n-1)!}\,\frac{g(x,t)}{g_t(x,t)}(x-t)^{n-1}.$$

The assumption that R_n can be expressed in the form (5) is not usually valid, of course, and we must modify our procedure somewhat.

Since we are concerned with $R_n(x, t)$ only for *fixed* $t=a$, the above line of reasoning suggests we write

$$R_n(x,a) = K(x)g(x,a), \tag{6}$$

where g is a given function (whose properties will be specified), and redefine the function F_x preceding (3) by

$$F_x(t) = f(x) - \left\{f(t) + \frac{f'(t)}{1!}(x-t) + \cdots + \frac{f^{(n-1)}(t)}{(n-1)!}(x-t)^{n-1} + K(x)g(x,t)\right\}. \tag{7}$$

F_x is no longer identically zero on $I \times I$, but $F_x(a)=0$ by (6), while

$$F_x(x) = -K(x)g(x,x).$$

This implies that $F_x(x)\equiv 0$ *if g is selected so that $g(x, x)\equiv 0$ on I.* (Of course $K(a)$ is not then defined by (6), but this is immaterial since $R_n(a, a)=0$.) F_x is continuous on I for each $x\in I$ *provided g is continuous* on $a\leqq t\leqq b$ for each $x\in I$. Finally, as in (3), we obtain

$$F_x'(t) = -\frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} - K(x)g_t(x, t), \qquad a < t < b, \tag{8}$$

assuming that g_t exists on $a<t<b$ for each $x\in I$. In particular, if $a<x\leqq b$, it follows from Rolle's theorem that $\exists X\in(a, x)$ such that

$$0 = F_x'(X) = -\frac{f^{(n)}(X)}{(n-1)!}(x-X)^{n-1} - K(x)g_t(x, X).$$

It now follows that *if $g_t(x, t)\neq 0$ for $a<x\leqq b$, $a<t<x$,* then

$$K(x) = -\frac{f^{(n)}(X)}{(n-1)!}\frac{(x-X)^{n-1}}{g_t(x, X)},$$

and

$$R_n(x, a) = -\frac{f^{(n)}(X)}{(n-1)!}\frac{g(x, a)}{g_t(x, X)}(x-X)^{n-1}, \qquad a < X < x. \tag{9}$$

We have proved the following

THEOREM. *If $f^{(n-1)}$ is continuous on $I=[a, b]$, and $f^{(n)}$ exists on $I^0=(a, b)$, then for each $x\in(a, b]$, f has the representation* (1), *where R_n is given by* (9), *g being any function satisfying the conditions*

(a) *g is continuous on $a\leqq t\leqq b$ for each $x\in I$;*
(b) *$g(x, x)\equiv 0$ on I;*
(c) *g_t exists on $a<t<b$ for each $x\in I$;*
(d) *$g_t(x, t)\neq 0$ for $a<x\leqq b$, $a<t<x$.*

Example 1. $g(x, t)\equiv h(x-t)$, where $h(t)$ is continuous for $t\geqq 0$, $h(0)=0$ and $h'(t)\neq 0$ for $t>0$. The remainder is given by

$$R_n(x, a) = \frac{f^{(n)}(X)}{(n-1)!}\frac{h(x-a)}{h'(x-X)}(x-X)^{n-1}, \qquad a < X < x. \tag{10}$$

Taking $h(t)=t^p$ for any $p>0$, we obtain *Schlömilch's remainder* ([2] p. 99)

$$R_n(x, a) = \frac{f^{(n)}(X)}{(n-1)!p}(x-X)^{n-p}(x-a)^p, \qquad a < X < x.$$

The choices $p=n$, $p=1$ give the remainders of Lagrange and Cauchy, respectively.

Example 2. $g(x, t) \equiv h(x) - h(t)$, where h is continuous on I, and $h' \neq 0$ on I_0. In this case the remainder is given by

$$R_n(x, a) = \frac{f^{(n)}(X)}{(n-1)!} \frac{h(x) - h(a)}{h'(X)} (x - X)^{n-1}, \qquad a < X < x. \tag{11}$$

Taking $h(t) = (t-a)^p$ for any $p > 0$, we obtain

$$R_n(x, a) = \frac{f^{(n)}(X)}{(n-1)!\,p} \frac{(x-a)^p}{(X-a)^{p-1}} (x - X)^{n-1}, \qquad a < X < x.$$

If $p = 1$ this is just the Cauchy remainder. The choice $p = n$ gives the remainder

$$R_n(x, a) = \frac{f^{(n)}(X)}{n!} \left(\frac{x - X}{X - a}\right)^{n-1} (x - a)^n.$$

Comparing the usual proofs (cf. [1] p. 136) of Taylor's formula with integral remainder (depending on successive integrations by parts) with that given above, there appears to be some pedagogical advantage to the latter: both this exact form of the remainder, and the special forms are obtained by essentially the same method.

References

1. W. Fulks, Advanced calculus, Wiley, New York, 1961.
2. W. Maak, An introduction to modern calculus, Holt, Rinehart and Winston, New York, 1963.

TAYLOR'S FORMULA WITH DERIVATIVE REMAINDER

ALFRED J. MARIA, late of Brooklyn College

In the usual derivation of Taylor's formula with derivative remainder there is used either an assumed form of the remainder or a generalized mean value theorem. In this note we shall derive Taylor's formula with a general form of derivative remainder as a straightforward application of the mean value theorem without any *a priori* assumed form of the remainder and also without use of generalized mean value theorems.

Let f be a function which together with its first n derivatives (finite) is continuous on the interval $a \leqq z \leqq x$, $a < x$ ($x \leqq z \leqq a$, $x < a$). It is further assumed that the $(n+1)$th derivative of f exists (finite) on the interval $a < z < x$ ($x < z < a$). We consider now the function F, defined for $a \leqq u \leqq x$, $a \leqq v \leqq x$ ($x \leqq u \leqq a$, $x \leqq v \leqq a$), whose value $F(u, v)$ at (u, v) is

$$f(u) - \sum_{j=0}^{n} \frac{(u - v)^j}{j!} f^{(j)}(v).$$

Essentially our problem is to find $F(x, a)$ in terms of a value of the $(n+1)$th derivative of f. To do this we join (x, a) to (x, x) by a suitably restricted curve $u = u(t)$, $v = v(t)$ and apply the mean value theorem to the function whose value

at t is $F(u(t), v(t))$. Differentiating this function we obtain

$$\frac{dF(u(t), v(t))}{dt} = f^{(1)}(u(t))u^{(1)}(t) - \sum_{j=1}^{n} \frac{(u(t) - v(t))^{j-1}}{(j-1)!} f^{(j)}(v(t)) \cdot (u^{(1)}(t) - v^{(1)}(t)) - \sum_{j=0}^{n} \frac{(u(t) - v(t))^{j}}{j!} f^{(j+1)}(v(t))v^{(1)}(t),$$

which after simplification gives

$$\frac{dF(u(t), v(t))}{dt} = f^{(1)}(u(t))u^{(1)}(t) - \frac{(u(t) - v(t))^{n}}{n!} f^{(n+1)}(v(t))v^{(1)}(t) - \sum_{j=1}^{n} \frac{(u(t) - v(t))^{j-1}}{(j-1)!} f^{(j)}(v(t))u^{(1)}(t).$$

We now choose the curve $u=x$, $v=v(t)$, where $v(0)=a$, $v(1)=x$ and $a<v(t)<x$ $(x<v(t)<a)$ for $0<t<1$. It is further assumed that v is continuous on the interval $0 \leqq t \leqq 1$ and is differentiable (finite) on the interval $0<t<1$. By the mean value theorem there is a number $\bar{t}$, $0<\bar{t}<1$, such that

$$F(x, x) - F(x, a) = -\frac{(x - v(\bar{t}))^{n}}{n!} f^{(n+1)}(v(\bar{t}))v^{(1)}(\bar{t}).$$

We therefore have, since $F(x, x)=0$,

$$f(x) = \sum_{j=0}^{n} \frac{(x - a)^{j}}{j!} f^{(j)}(a) + R_{n+1},$$

where we have set $R_{n+1} = \{(x-v(\bar{t}))^{n}/n!\}f^{(n+1)}(v(\bar{t}))v^{(1)}(\bar{t})$. This is the desired result.

It is to be observed that if $v^{(1)}(t) \neq 0$ ($v^{(1)}(t)$ finite) for $0<t<1$, then it is sufficient to assume that the $(n+1)$th derivative of f exists (finite or infinite) on the interval $a<z<x$ $(x<z<a)$. If we take $v(t)=x-(x-a)(1-t)^{1/p}$, $p>0$, we shall have

$$R_{n+1} = \frac{(x - a)^{n+1}}{n!} f^{(n+1)}(x - (x - a)(1 - \bar{t})^{1/p}) \frac{1}{p} (1 - \bar{t})^{(n+1)/p-1},$$

and on setting $(1-\bar{t})^{1/p} = 1-\theta$ we obtain

$$R_{n+1} = \frac{(x - a)^{n+1}}{n!} f^{(n+1)}(a + \theta(x - a)) \frac{1}{p} (1 - \theta)^{n+1-p},$$

which is Schlömilch's form of the remainder.

A NEW PROOF OF TWO SUMMATION FORMULAS

C. W. BARNES, University of Mississippi

Heath [1], p. 114, states that certain proportions in Archimedes are equivalent to the formulas

$$\sin\frac{\pi}{2n}+\sin\frac{2\pi}{2n}+\cdots+\sin\,(2n-1)\,\frac{\pi}{2n}=\cot\frac{\pi}{4n}$$

and

$$\frac{2\left(\sin\dfrac{x}{n}+\sin\dfrac{2x}{n}+\cdots+\sin\,(n-1)\,\dfrac{x}{n}\right)+\sin x}{1-\cos x}=\cot\frac{x}{2n}\cdot$$

Heath also pointed out that Archimedes' work gives in fact a summation of the series $\sin x+\sin 2x+\cdots+\sin\,(n-1)x$. Reference is made to Loria, *Il periodo aureo della geometria greca.*

A particularly simple proof using complex numbers is given by Loney ([2], pp. 115–116) for the more general identities

$$(1)\qquad 1+\sum_{k=1}^{n-1}t^k\cos kx=\frac{1-t\cos x-t^n\cos nx+t^{n+1}\cos\,(n-1)x}{1-t\cos x+t^2}$$

and

$$(2)\qquad \sum_{k=1}^{n-1}t^k\sin kx=\frac{t\sin x-t^n\sin nx+t^{n+1}\sin\,(n-1)x}{1-2t\cos x+t^2}.$$

It is unfortunate that these formulas are so little known, as the writer has found that they are useful in beginning calculus classes in the evaluation of such integrals as $\int_0^{\pi/2}\sin x\,dx$ by appealing to the definition of a Riemann integral. The case $t=1$ is used for this, where $x\neq 2k\pi$ for k an integer.

For example, if the existence of $\int_0^{\pi/2}\sin x\,dx$ has been discussed and if $\lim_{x\to 0}(\sin x/x)=1$ has been proved, we divide the interval $[0,\pi/2]$ into n equal subintervals, evaluate $\sin x$ at the right endpoint of each subinterval and form the approximating sum

$$S_n=\sum_{k=1}^{n}\frac{\pi}{2n}\sin\frac{k\pi}{2n}\cdot$$

By (2), this sum may be written

$$S_n=\frac{\pi}{2n}\left\{\frac{\sin\dfrac{\pi}{2n}-\sin\dfrac{(n+1)\pi}{2n}+\sin\dfrac{\pi}{2}}{2-2\cos\dfrac{\pi}{2n}}\right\}$$

or when simplified,

$$S_n = \frac{\pi}{4n} + \frac{[(1 + \cos(\pi/2n))]/2}{\sin(\pi/2n)/(\pi/2n)} \cdot$$

Letting n increase, we have

$$\int_0^{\pi/2} \sin x \, dx = \lim_{n\to\infty} S_n = 1.$$

Recently the writer was able to give a new proof of (1) and (2) by using methods of elementary matrix algebra. The proof is an interesting application of some of the routine theorems in a first course in matrices.

Consider the matrix

$$A = \begin{bmatrix} t\cos x & -t\sin x \\ t\sin x & t\cos x \end{bmatrix}, \tag{3}$$

where x and t are real. We verify by induction that for every positive integer n

$$A^n = \begin{bmatrix} t^n\cos nx & -t^n\sin nx \\ t^n\sin nx & t^n\cos nx \end{bmatrix}.$$

The following theorem is proved in Mirsky [3] page 332.

THEOREM 1. *If the moduli of all characteristic roots of the matrix A are less than* 1, *then $I-A$ is nonsingular and the series $I+A+A^2+A^3+\cdots$ converges to $(I-A)^{-1}$.*

Moreover, we have

THEOREM 2. *If $I-A$ is nonsingular then $I+A+A^2+A^3+\cdots+A^{n-1} =(I-A^n)(I-A)^{-1}$.*

We note that the characteristic roots of (3) are te^{ix} and te^{-ix}. Hence Theorem 1 may be applied if we require $|t|<1$.

Since

$$(I - A)^{-1} = \frac{1}{1 - 2t\cos x + t^2}\begin{bmatrix} 1 - t\cos x & -t\sin x \\ t\sin x & 1 - t\cos x \end{bmatrix},$$

whenever $|t|\neq 1$, and $x\neq 2k\pi$ if $|t|=1$, Theorem 2 gives

$$\begin{bmatrix} 1 + \sum_{k=1}^{n-1} t^k\cos kx & -\sum_{k=1}^{n-1} t^k\sin kx \\ \sum_{k=1}^{n-1} t^k\sin kx & 1 + \sum_{k=1}^{n-1} t^k\cos kx \end{bmatrix}$$

$$=\begin{bmatrix} \dfrac{1-t\cos x-t^n\cos nx+t^{n+1}\cos(n-1)x}{1-2t\cos x+t^2} & \dfrac{-t\sin x+t^n\sin nx-t^{n+1}\sin(n-1)x}{1-2t\cos x+t^2} \\ \dfrac{t\sin x-t^n\sin nx+t^{n+1}\sin(n-1)x}{1-2t\cos x+t^2} & \dfrac{1-t\cos x-t^n\cos nx+t^{n+1}\cos(n-1)x}{1-2t\cos x+t^2} \end{bmatrix}$$

after simplifying the elements of the product on the right. Equating corresponding elements of these matrices we obtain (1) and (2).

If we now suppose $|t|<1$ we may apply Theorem 1 to obtain

$$\begin{bmatrix} 1+\sum_{k=1}^{\infty} t^k\cos kx & -\sum_{k=1}^{\infty} t^k\sin kx \\ \sum_{k=1}^{\infty} t^k\sin kx & 1+\sum_{k=1}^{\infty} t^k\cos kx \end{bmatrix} = \begin{bmatrix} \dfrac{1-t\cos x}{1-2t\cos x+t^2} & \dfrac{-t\sin x}{1-2t\cos x+t^2} \\ \dfrac{t\sin x}{1-2t\cos x+t^2} & \dfrac{1-t\cos x}{1-2t\cos x+t^2} \end{bmatrix}.$$

It follows that

$$1+\sum_{k=1}^{\infty} t^k\cos kx = \frac{1-t\cos x}{1-2t\cos x+t^2}$$

and

$$\sum_{k=1}^{\infty} t^k\sin kx = \frac{t\sin x}{1-2t\cos x+t^2}$$

as found in [2], p. 116.

References

1. T. L. Heath, The Works of Archimedes, Dover, New York, 1953.
2. S. L. Loney, Plane trigonometry, part 2, Cambridge University Press, Cambridge, 1952.
3. L. Mirsky, An Introduction to Linear Algebra, Oxford University Press, London, 1955.

PICTURES ABOUT 0-DERIVATIVES

BURROWES HUNT, Reed College

There is some interest in proving as directly as possible the theorem that a function with 0-derivative on an interval must be constant on the interval. [See for example: Mary Powderly, this MONTHLY, 70 (1963) 544.]

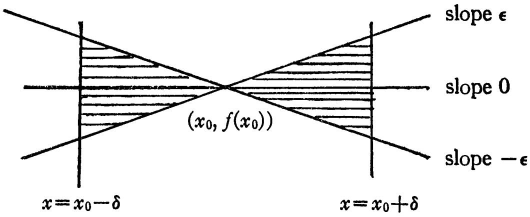
Suppose that the function g is such that $g'=0$ on an interval J, but g is not constant on J. There are a and b in J with $a<b$ and $g(a)\neq g(b)$. Define f, for x in the domain of g, by $f(x)=(g(x)-g(a))^2$. Then $f'=0$ on J, $f(a)=0$, and $f(b)>0$.

Picture A is about the graph of a function f such that $f'(x_0)=0$, for some x_0. For each $\epsilon>0$ there is $\delta>0$ such that the graph of f is in the interior of the

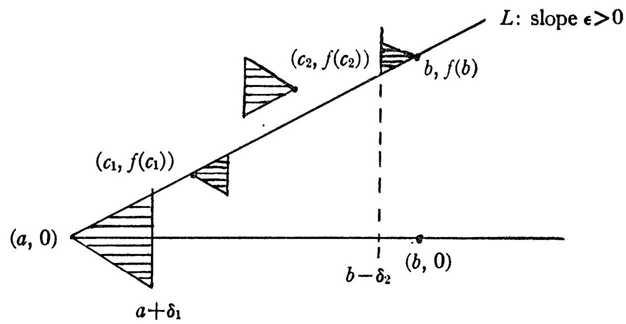
shaded triangles over the intervals $(x_0-\delta,\ x_0)$ and $(x_0,\ x_0+\delta)$. In short, A is a picture about:

$$\left|\frac{f(x) - f(x_0)}{x - x_0} - 0\right| < \epsilon.$$

With f as in the second paragraph let L, with equation $y=L(x)$, be the line through $(a, 0)$ and $(b, f(b))$. L has positive slope: call it ϵ.



PICTURE A



PICTURE B

The right half of Picture A, applied in Picture B with $x_0=a$, shows that over an interval $(a,\ a+\delta_1)$, the graph of f is under L. The left half of Picture A, applied with $x_0=b$, shows that over an interval $(b-\delta_2,\ b)$, the graph of f is above L.

Let S be the set of x in (a, b) such that $f(x) \leqq L(x)$. Then S is nonempty because $a+\frac{1}{2}\delta_1$ is in S, and has upper bounds: e.g. $b-\delta_2$.

Picture B makes it clear that if c_1 is in (a, b) and $f(c_1) \leqq L(c_1)$ then c_1 is not an upper bound for S. Also it is clear that if $f(c_2) \nleqq L(c_2)$ and c_2 is an upper bound of S then c_2 is not the least upper bound.

The existence of a nonconstant function with derivative 0 on an interval contradicts the least upper bound property.

AN ELEMENTARY CALCULUS COMPUTATION WITHOUT VARIABLES

ALBERT WILANSKY, Lehigh University

1. A recent note [1] concerned solutions of the functional equation $Df[g(x)] = g(x)$. Without making any mathematical contributions, we point out that the notations of calculus "without variables," as proposed in [2], make the computations simpler and clearer and avoid the use of *ad hoc* designations. The calculations of [1] are carried out in section 3. The two examples given are the same as those in [1]. Section 2 is devoted to an exposition of the notation of [2], modified by the present author without the knowledge or consent of Professor Menger. In the author's opinion this notation is not as well known as it deserves to be.

2. Numbers are denoted by a, b, c, x, y; functions by f, g, h. Also, $a\cdot f$, $f\cdot g$, $f+g$, fg are those functions whose values at x are

$$a\cdot fx,\ (fx)\cdot(gx),\ fx + gx,\ f[gx]$$

respectively. Composition of functions is given first priority, thus

$$fg\cdot hx = (fgx)\cdot hx = f[g(x)]\cdot hx.$$

The function j is defined by $jx = x$. Obvious definitions lead to such equations as $(j^{1/2}+j^4)f = \sqrt{f}+f^4$, $(1/j)f = 1/f$. Constant functions are denoted by underlines, thus $\underline{2}x = 2$, and $(j^2+\underline{2})x = x^2+2$; j^2+2 is meaningless; $x^2+2x = (j^2+2j)x$.

Df is the derivative of f and D reaches only to the succeeding letter, for example $Dj^3 \sin = 3j^2 \sin = 3 \sin^2$. The chain rule reads $D(fg) = Dfg\cdot Dg$, for example $D \sin^3 = D(j^3 \sin) = Dj^3 \sin\cdot D \sin = 3j^2 \sin\cdot\cos = 3 \sin^2\cdot\cos$.

(The next notation is violently objected to on p. 155 of [2].)

$\int f$ is (the) (an) indefinite integral of f and $\int$ reaches only to the succeeding letter, for example $\int j^2 \cos = (j^3/3) \cos = \cos^3/3$. The inverse chain rule reads $\int(fg\cdot Dg) = \int fg$. For example, to evaluate the traditionally written $\int(\cos x^3)3x^2\,dx$ we have

$$\int(\cos j^3\cdot 3j^2) = \int(\cos j^3\cdot Dj^3) = \int\cos j^3 = \sin j^3 \text{ or, for } \int(2x\,dx)/(1+x^2)$$

write $\int[1/j(1+j^2)\cdot D(\underline{1}+j^2)] = \int 1/j(\underline{1}+j^2) = \log\ (\underline{1}+j^2)$.

3. To solve $D(fg) = g$.

CASE I: g is given. Then $fg = \int g$ so $f = \int g\ g^{-1}$.

Example 1. $g = \log$. Then $f = \int\log\exp = [j\cdot(\log-\underline{1})]\exp = (j\exp)\cdot(\log-\underline{1})\exp = \exp\cdot(j-\underline{1})$ or $f(x) = e^x(x-1)$.

CASE II: f is given. Then $g = D(fg) = Dfg\cdot Dg$. Hence

$$\underline{1} = \frac{Dfg}{g}\cdot Dg = \left(\frac{Df}{j}\right)g\cdot Dg \quad\text{and so}\quad j = \int\underline{1} = \int\frac{Df}{j}\,g,$$

and so, finally

$$g = \left(\int \frac{Df}{j}\right)^{-1}.$$

Example 2. $f = cj^n$. Then

$$g = \left(\int ncj^{n-2}\right)^{-1} = \left(\frac{nc}{n-1} j^{n-1}\right)^{-1} = j^{1/(n-1)}\left(\frac{n-1}{nc} j\right)$$

by Lemma 1, below. Thus

$$g(x) = \left[\frac{(n-1)x}{nc}\right]^{1/(n-1)}.$$

LEMMA 1. $(aj^n)^{-1} = j^{1/n}(j/a)$.

For $(aj^n)(j^{1/n})(j/a) = aj(j/a) = j$.

References

1. D. W. Hight and A. G. Haddock, $D_x f[g(x)] = g(x)$, this MONTHLY, 71 (1964) 1034–1035.
2. K. Menger, Calculus, A modern approach, Ginn and Company, Boston, 1955.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and
JOHN A. BROWN, University of Delaware

All material for this department should be sent to John R. Mayor 1515 Massachusetts Avenue, N. W., Washington, D. C. 20005.

TEACHING EXPERIENCE FOR PROSPECTIVE TEACHERS OF MATHEMATICS

V. D. TURNER, Mankato State College

This is a report on an attempt to provide teaching experience for prospective teachers of mathematics at Mankato State College in addition to their regular student teaching program.

For the past year, a two-quarter hour special problems course called "Teaching Experience" has been offered to enable junior or senior mathematics majors to work closely with staff members. These students are called special assistants and, thus far, have been assigned to staff members teaching the general education course. This course meets four days per week covering the idea of sets, structure of the number system, number bases, logic, and elementary algebraic concepts. Classes generally have an enrollment of 70–80 students.

The special assistant was invited to attend lectures as deemed necessary to

acquaint himself with the course content, method used by the instructor, kinds of difficulties experienced by the class, and so on. The class met with the regular instructor four days per week, and on the fifth day a help session was conducted by the special assistant. During the quarter, the special assistant also made out an hour examination under the direction of the instructor, assisted in administering the examination, graded the examination, and prepared an item analysis of the test.

The response from special assistants participating to date has been very enthusiastic. Typical comments are:

"I feel that I have gained considerable confidence in appearing before a group and am sure that I want to teach."

"I appreciate the opportunity to work closely with what goes on behind the scenes. I think I would like college teaching."

"I would say the experience I gained was invaluable, so much so in fact that these two credits could well be the most useful a future mathematics teacher could take."

The staff feels that the increased personal attention is of benefit in the large class sections, and student acceptance is evidenced by an average of 25–30 at each help session.

In addition, the staff has found this program to be helpful in that it enables them to get better acquainted with mathematics majors.

PURE MATHEMATICS PUBLICATIONS: 1939–1958

L. R. Walum, Denison University

This paper reports some statistical and descriptive results concerning publication in pure mathematics from 1939 to 1958. Pure mathematics was delimited to the five specialties of algebra, geometry, analysis, topology and number theory. A systematic random sample of approximately 9 per cent of the publications so classified in the *Mathematical Reviews* was taken.

The total sample numbered 5114 which estimates a population of 57,000 articles for those nineteen years. Although the number of publications increases steadily, two major troughs are noteworthy. The first occurs during the middle of World War II; the second during the Korean War.

Forty-one nations are included as publishing loci for mathematics. Of these, however, seven account for approximately 70 per cent of the total frequency. These are the United States (21 per cent), the U.S.S.R. (14 per cent), Italy (9 per cent), France (8 per cent), Japan (7 per cent), Germany (7 per cent) and England (5 per cent). Since the rank order of nations is maintained, year by year, through time, the ordering is not the result of publication during any one period, including the present one.

When national publication is standardized by population size, the leading nations are Switzerland, Belgium, Norway, Israel, Netherlands and Hungary in that order.

Although there are 26 different languages represented, including Gaelic and Esperanto, five major ones account for nearly 92 per cent of the total sample. These are English (43 per cent), French (16 per cent), Russian (13 per cent), German (11 per cent), and Italian (8 per cent). All of the Scandinavian languages combined account for only .5 per cent of the total and all of the Asian languages for only .6 per cent.

The mathematical specialties maintain the same relative positions through time. In order of preponderance, they are analysis (50 per cent), geometry (20 per cent), algebra (14 per cent), number theory (9 per cent), and topology (7 per cent).

Over 92 per cent of the publications are of single authorship. However, there is a significant trend towards increasing joint authorship, particularly in the United States and Japan. All of the mathematical specialties are about equally prone to single author papers.

From "Pure Mathematics: A Study in the Sociology of Knowledge," (Unpub. Ph.D. thesis, Univ. of Colo., 1963). This research was supported by a Public Health Service Predoctoral Fellowship (MPM-12,836) from the National Institute of Health. I thank Edward Rose for guiding this project, Robert C. Hanson, Alex Garber, Sarvadaman Chowla, and Chandler Davis for their suggestions and advice.

BRIEF COMMENT

In response to requests from readers of the MONTHLY a new feature has been added to Mathematical Education Notes. Professor John D. Baum, of Oberlin College, has kindly agreed to review articles in other publications which are believed to be of interest to readers of the MONTHLY.

"But is the teacher also a citizen?" ALVIN M. WEINBERG, *Science*, August 6, No. 3684, 149(1965) 601–606.

"Our society is 'mission-oriented.' Its mission is resolution of problems arising from social, technical, and psychological conflicts and pressures. The university by contrast is 'discipline-oriented.' Its viewpoint is the sum of the viewpoints of the separate, traditional disciplines that constitute it. The problems it deals with are, by and large, problems generated and solved within the disciplines themselves. In society the nonspecialist and synthesizer is king. In the university the specialist and analyst is king."

This article, dealing as it does with "mission-oriented" society vis-à-vis "the discipline-oriented" university, points up the conflict between the two. Choosing many examples from the field of the teaching of mathematics in the schools—particularly emphasizing the recent radical reforms and the extent to which they have been dictated by the university—the author questions whether such changes serve society's interests best. He argues that "the discipline-oriented university encourages the rise of purism and specialization and the denial of scholarship and application" and that "these trends in the universities are affecting our elementary curricula; are giving us poorer people to get on with

the applied work of the day; are substituting research for action; and are tending to impose the scientific values of the fragmented university upon society."

An editorial in the New York Times for August 10, 1965, page 28, discusses the article.

Experiments in the teaching of sixth form mathematics to nonspecialists, Pamphlet No. 1 of a series prepared for the Mathematical Association, London, 1965.

"Ideally, every adult should understand the simple mathematics of citizenship and the national budget, and should be able to distinguish between the use and misuse of statistics. Certainly no sixth former should be afraid of decimal points, or strings of noughts in large numbers, or simple proportion. It would be difficult, perhaps impossible, to devise a course which would attract all types of pupils. The object of this pamphlet is to describe some of the work which is going on. While the accounts may suggest possibilities for various kinds of pupil, it must be emphasised that they are not intended to be in any way final, authoritative or exclusive."

"We cannot state too clearly nor too often that the object of any non-specialist course is to encourage critical thinking, and not manipulative technique based upon rule of thumb. We want our pupils to have understanding, to be able to analyse a variety of situations and pick out those features which are amenable to mathematical treatment. We want them to be able to solve interesting and significant problems by bringing to bear some depth of understanding, not by applying a standard method."

With the above statement of philosophy as background and with some further discussion of this philosophy, there are described three programs which have been carried on in recent years at various British schools. There is also a short booklist.

The information revolution, Section 11, *The New York Times,* May 23, 1965.

This advertising supplement was sponsored by The International Federation for Information Processing and participating advertisers. It contains a sequence of articles dealing with information storage and retrieval via computers. Of particular interest, perhaps, is the article by R. A. Buckingham of the University of London on *Education in Information Processing* in which it is asserted that "training in how to use a computer must begin far below the university level and must be provided to a far broader spectrum of the population than just technical people. For there is no doubt that the computer during the next few decades will evolve from a mysterious presence used by a relatively few people into a commonplace tool used, in some way or other by almost everyone."

Copies of the supplement can be obtained from the American Federation of Information Processing Societies, 211 East 43rd St., New York, N. Y. 10017, at 50 cents each.

A second report on the teaching of mechanics in schools, A report prepared for the Mathematical Association, London, 1965.

Besides detailed course descriptions and a brief history of mechanics up to the time of Newton, the report presents a philosophy of the teaching of mechanics in schools. The philosophy is best stated in the words of the report: "We have no hesitation in commending mechanics as a part of a secondary school mathematics course, whatever the age or grade of the pupils and whether they be in the early years in a secondary modern school or preparing for university entrance. It is a subject which, more than most, can give reality to mathematics by showing how it may be applied. In the hands of a skilled teacher it can be used to stimulate pupils to face and overcome the mathematical difficulties that present themselves."

Emerging twelfth-grade mathematics programs, LAUREN G. WOODBY, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 1965 (Catalog No. FS 5.229:29060) v+40 pgs, 20 cents.

"This report describes the fourth- and fifth-year courses in mathematics offered during 1962–63 by 66 selected high schools in 20 States. The report is an intensive study of some current practices that appear to be promising, rather than an extensive study designed to give a composite view of the situation in all the schools in the Nation."

The conclusions of the report are about what one would expect: good teachers teach good programs; there is no agreement on what the best course for this level is; concern as to what to do with calculus continues; the following conclusion, however, is not what this reviewer would have predicted: "There is little acceptance of a course in probability and statistics as the fourth or fifth year mathematics offering in the college-preparatory program." Of the recommendations, the following is worth mentioning: "Courses in calculus and analytic geometry of the warmup variety are not recommended." Instead either a full year of calculus (perhaps, advanced placement style) or courses in probability and statistics, linear algebra, or analytic geometry should be considered.

An experiment in the teaching of "Modern Mathematics" in schools, ROBIN J. WILSON, *The Mathematical Gazette*, 49 (1965) 22–33.

This article is an account of an experiment in teaching two forms in a British Public School (University College School) a two week course covering several topics in 'modern mathematics.' The boys, 46 in number, were of two ages, 12 year olds and 14 year olds. The topics covered were matrices, sets, number systems, and groups. The most interesting aspect of the article is perhaps the portion which deals with the reaction of the students to the material. Few of them seem to have found it very difficult, and many found one topic or another quite stimulating. The following statement on the part of one student is worth repeating: "The present syllabus is terribly uninteresting, and we have

to wait until we are at University to do the beautiful and interesting calculations. If children did so-called 'Modern Maths' at an early age, we would be far more advanced in our ideas. Doing 2000 year old Maths is just as silly as doing 2000 year old Physics e.g. Earth, Fire, Air and Water theory."

Stage A Topology in the Main School, A. R. TAMMADGE, *The Mathematical Gazette*, 48 (1964) 365–372.

Despite the title of this article, it deals largely with graph theory. The article describes the material presented to a class of second form (12+ year old) boys in 1964. The assertion is made that the class was not particularly bright. The author makes an association of a square matrix with each linear graph; thus the article, besides providing an interesting example of enrichment material for secondary school use, also provides a strong motivation for the introduction of matrices and their properties.

The AAAS Science Book List for Young Adults, HILARY J. DEASON, (Editor), AAAS Miscellaneous Publication No. 64-11, Washington, D. C., 1964.

The books listed in this book list "are intended primarily for collateral reading and reference by students in the ninth to twelfth grades, for students in equivalent forms in private preparatory schools, for first and second year college students, and for their instructors. They are also appropriate for the interested adult reader." Of the 1376 titles listed some ten per cent are in the field of mathematics. The book list is annotated with brief comments on each of the titles; these comments should prove quite useful for someone seeking reference material in a particular area and at a particular level. The book list is available from the American Association for the Advancement of Science, 1515 Massachusetts Ave., N.W., Washington, D.C. 20005 for $2.50 in the paperback edition, and $3.50 in cloth.

The Comprehensive College Tests, Educational Testing Service, Princeton, 1964.

This ten page descriptive brochure announces the issuance by ETS of a sequence of examinations at the college level. These examinations have been provided, since "in recent years, colleges throughout the country have expressed a need for college-level tests that would provide measures of undergraduate achievement that could be used for a variety of purposes." The tests offered fall into two categories, the first of which is "the general examinations [which] are intended to provide a comprehensive measure of undergraduate achievement in the five basic areas of the liberal arts—English Composition, Humanities, Mathematics, Natural Sciences, and Social Sciences-History." The other category is that of subject examinations which are "end-of-course examinations that are developed for certain widely-taught undergraduate courses most frequently taken in the first two years of college." Among the subject examinations is one in introductory calculus. "Each Subject Examination consists of a 90-minute,

multiple-choice test with an optional essay section. The essay test requires an additional 90 minutes of testing time." Further information concerning the program and the tests in the program may be obtained by writing to Director, Comprehensive College Tests, Educational Testing Service, Princeton, N. J. 08540.

Wordless programing of mathematical concepts, M. DANIEL SMITH, *Programed Instruction,* June 1965 (Vol. IV, No. 9).

The above programs were written in response to "difficulties in studying abstract mathematics, some of which seemed to stem from a need for verbal mediators in areas where such mediators were ineffective." There seemed to be to the author "certain advantages to learning sequences which did not use words, and which therefore in some sense shaped the ability to employ nonverbal mediating processes." The "non-wordal" programs developed were given to five fifth-grade students in the academic year 1962 and covered such topics as vectors interpreted as operators, vector addition, addition of directed numbers, inverse, subtraction of vectors and of directed numbers, nonrectangular coordinate systems, and addition and subtraction of ordered sets (sic) and matrices. "While the students had little trouble learning the material, they did have some trouble explaining what they were learning to their parents, since they had no words for it." Further work in this area seems indicated, for one wonders how successfully a concept is grasped if it is impossible to communicate it adequately.

The Ubiquitous TA, *Time Magazine,* June 4, 1965 (Vol. 85, No. 23) 49–50.

This article, which describes the position and function of the teaching assistant in the American University today, gives a fair picture of the sorts of problems which are faced by mathematics departments—as well as many other departments—in many American universities. There are some interesting statistics on the number of teaching assistants at various schools, as well as on salary levels for teaching assistants.

Review of **Goals for School Mathematics: The Report of the Cambridge Conference on School Mathematics,** Boston, Houghton Mifflin Company, 1963, by MARSHALL H. STONE, in *The Mathematics Teacher,* 58 (1965) 353–360.

"I am reluctant to believe that the Cambridge Report represents the best thinking of which we in the United States are collectively capable in the field of mathematical education. The goals proposed by the Cambridge Report fall short of those already formulated and already on the way to realization in Europe. The merit of the Cambridge Report is not to be found in its substance —in that respect it is, quite bluntly, extremely disappointing—but rather in its willingness to challenge the extent of our current achievements in the field of mathematical education and to demand a thorough and uncompromising revision of the entire school mathematics curriculum." The above quotes from the review give one the tone of the criticism—criticism of the most constructive

kind, for besides making explicit the objections to the findings of the Report, and the reasons for these objections, alternative solutions are proposed.

This is an important and valuable review. It deserves the attention of anyone who is concerned with and involved in mathematical education.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE, Bloomfield College

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, Bloomfield College, Bloomfield, N. J. 07003. *Proposers of Problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.*

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to A. E. Livingston, University of Alberta, Edmonton, Alberta, Canada. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before May 31, 1966.

E 1845. *Proposed by R. S. Nickerson, Hanscom Field, Bedford, Mass.*

Consider a sequence of integers in which: (a) every integer, i, $0<i\leqq n$, occurs exactly twice, and (b) the second occurrence of an integer, i, follows its first occurrence by exactly i integers. For example, for $n=5$, the sequence 3, 5, 2, 3, 2, 4, 5, 1, 1, 4 satisfies these conditions.

For arbitrary n, (1) state a rule determining whether or not at least one such sequence exists, (2) describe a procedure for generating such a sequence if it exists, and (3) state a rule for determining how many such sequences exist.

E 1846. *Proposed by M. N. S. Swamy, University of Saskatchewan*

Show that the Fibonacci polynomial defined by

$$f_n(x) = x\cdot f_{n-1}(x) + f_{n-2}(x), \qquad f_1 = 1, \qquad f_2 = x,$$

satisfies the inequality $f_n^2(x)\leqq(x^2+1)^2\cdot(x^2+2)^{n-3}$, $n>2$.

E 1847. *Proposed by J. I. Nassar, Socony Mobil Oil Company, Princeton, N. J.*

Let h_i, r_i $(i=1, 2, 3)$ be the altitudes and exradii respectively of a triangle,

and let α_1, α_2, α_3 be any nonnegative numbers. Prove that

$$\sum h_i^{\alpha_1} h_j^{\alpha_2} h_k^{\alpha_3} \geqq \sum r_i^{\alpha_1} r_j^{\alpha_2} r_k^{\alpha_3},$$

where the summation is taken over all permutations of 1, 2, 3. The equality holds if and only if the triangle is equilateral. Here x^α is taken to be the positive number x^α for $x>0$, $\alpha \geqq 0$. This is an extension of E 1675 [1965, 187].

E 1848. *Proposed by J. F. Traub, Bell Telephone Laboratories, Murray Hill, N. J.*

Let $P(t)$ be a polynomial of degree n with distinct zeros $\rho_1, \rho_2, \cdots, \rho_n$. We know from interpolation that

$$P'(t) = \sum_{i=1}^{n} P'(\rho_i) L_i(t),$$

where $L_i(t)$ is the Lagrange polynomial

$$L_i(t) = \frac{P(t)}{P'(\rho_i)(t-\rho_i)}.$$

Prove also that $P'(t) = \sum_{i=1}^{n} P'(\rho_i) L_i^2(t)$.

E 1849. *Proposed by George Purdy, The University, Reading, England*

Let $f_0(x) = \cosh x$, $f_1(x) = \sinh x$. Then for $r = 0$, 1 and $m \geqq 1$,

$$f_r(x_1 + x_2 + \cdots + x_m) = \sum f_{s_1}(x_1) \cdots f_{s_m}(x_m)$$

identically, where the summation is taken over all choices of $s_1, \cdots, s_m$, each being 0 or 1, with $s_1 + \cdots + s_m \equiv r \pmod 2$. A proof is desired which does not employ mathematical induction.

E 1850. *Proposed by George Purdy, The University, Reading, England*

Prove the following formula in which $\sigma(n)$ is the sum of the divisors of the positive integer n, and $[x+1]$ denotes the integral part of $x+1$.

$$\sigma(n) = \sum_{m=1}^{n} \int_0^m \cos\left(\frac{2\pi n[x+1]}{m}\right) dx.$$

E 1851. *Proposed by W. E. Patten, South Boston, Va.*

In problem E 1704 [1965, 669] it was found that if an $m \times n$ rectangle is partitioned into mn unit squares in the natural way, then the number of these which meet a diagonal of the rectangle is $m+n-(m, n)$, where (m, n) is the g.c.d. of the positive integers m and n. Given the positive integer k, for how many rectangles does a diagonal meet precisely k unit squares?

E 1852. *Proposed by Stephen Haines and L. F. Leetch, Bowling Green State University, Ohio*

If $F^*(x) = \lim_{h\to 0}\{F(x+h)-F(x-h)\}/2h$, prove or disprove the following statement: If F and F^* are continuous at x, then F' exists at x and $F'(x)=F^*(x)$.

E 1853. *Proposed by D. R. Breach, University of Toronto*

A format for a crossword puzzle is formed by blacking in certain cells of a square of n^2 cells. In any row or column call the blank gaps between the black cells words. (With this definition a word might contain one cell only.) In a well-designed format any two words are connectable, that is, a rook may move from a cell of one to a cell of the other without passing through any black cells. How many well-designed formats are there?

E 1854. *Proposed by Rolf Schneider, the University, Frankfurt am Main*

Let p be an odd prime and k any positive integer. Show that

$$\sum_{i=0}^{(p^k-1)/2} (-1)^i \binom{2i}{i} \equiv 5^{k(p-1)/2} \pmod{p},$$

$$\sum_{i=0}^{(p^k-1)/2} \binom{2i}{i} \equiv (-3)^{k(p-1)/2} \pmod{p}.$$

SOLUTIONS OF ELEMENTARY PROBLEMS

Composite Components for Vector-valued Integral Polynomials

E 1748 [1965, 74]. *Proposed by Michael Gemignani, University of Notre Dame*

Let $S=\{P_1(x), \cdots, P_n(x)\}$ be any finite set of nonconstant polynomials with positive integer coefficients. Show that for some natural number m, the n integers $P_1(m), \cdots, P_n(m)$ are all composite.

Solution by William H. Bonney, New Mexico State University. Suppose merely that $P_1(x), \cdots, P_n(x)$ are nonconstant polynomials with integral coefficients and positive leading coefficient. (The latter requirement is no restriction since $P_1(m), \cdots, P_n(m)$ are all composite if and only if $\pm P_1(m), \cdots, \pm P_n(m)$ are). For some integer a, the polynomials $P_1(x), \cdots, P_n(x)$ are greater than or equal to 2 and strictly increasing for $x \geqq a$. Let a be so given. For each j among $1, \cdots, n$, the Taylor expansion of P_j about a has integral coefficients and constant term $P_j(a) \geqq 2$, and $P_j(x) > P_j(a)$ for $x > a$. Let M be the l.c.m. of $P_1(a), \cdots, P_n(a)$. For each positive integer k and each j among $1, \cdots, n$, we have $P_j(a) \mid P_j(a+kM)$ and $2 < P_j(a) < P_j(a+kM)$. Hence $P_1(a+kM), \cdots, P_n(a+kM)$ are composite positive integers for each positive integer k.

Also solved by J. C. Abad, R. G. Albert, D. P. Ambrose (Basutoland), J. T. Anderson, M. A. Bershad, Walter Bluger, J. A. Burslem, Leonard Carlitz, Mannis Charosh, Allan Chuck and Peter Goldstein (jointly), C. C. Clever, David Cohen, Daniel Cohen, J. E. Connett, R. B. Eggleton (Australia), E. W. Ewing, G. F. Feissner, N. J. Fine, W. D. Fryer, Philip Fung, P. K. Garlick, R. W. Gilmer, Jr., A. A. Gioia and A. M. Vaidya (jointly), Michael Goldberg, Jerry Goodman, H. S. Hahn, D. M. Hancasky, Ned Harrell, Susan H. Harris, B. A. Hausmann, G. A. Heuer, Ray Jurgensen, M. S. Klamkin, H. L. Lahmann (Germany), C. C. Lindner, Robert Maas, Andrzej Makowski (Poland), D. C. B. Marsh, Charles McCracken, M. G. Murdeshwar, Harsh Pittie, G. B. Purdy (England), Simeon Reich (Israel), Robin Sibson (England), L. W. Stern, G. C. Thompson, A. E. Tong, Guy Torchinelli, B. R. Toskey, C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), Emanuel Vegh, Julius Vogel, John Waddington, Lenard Weinstein, William Wernick, and the proposer.

A Bound for the Average of the Divisors of n

E 1749 [1965, 75]. *Proposed by Stanton Philipp, Long Beach State College, California*

Prove if $n>1$ then $\sigma(n)/\tau(n)\leqq 3n/4$, where $\tau(n)$ is the number of divisors of n, and $\sigma(n)$ is the sum of the divisors of n.

I. *Solution by Allan Chuck and Peter Goldstein, Students, Lincoln High School, San Francisco.* Call n/d the complement of the divisor d of n. n is the only divisor of n which is greater than or equal to $3n/4$, its complement is 1, and their arithmetic mean is $(n+1)/2$. If n is composite, all other divisor-complement pairs have arithmetic mean less than $3n/4$, since both divisor and complement are less than $3n/4$. But $(n+1)/2\leqq 3n/4$, with equality only for $n=2$. Thus the arithmetic mean of all the divisors of n is less than or equal to $3n/4$. Finally, the arithmetic mean is simply $\sigma(n)/\tau(n)$.

II. *Solution by E. S. Langford, U. S. Naval Postgraduate School, Monterey, California.* Note that $\sigma(n)/\tau(n)$ is the "average size" of a divisor of n. Consider then $f(d)=(d+n/d)/2$ for $1\leqq d\leqq n$. Since $f''(d)=n/d^3>0$, f is concave and, hence, has the maximum $(n+1)/2$. Since $(n-1)^2>0$, we see also that $n^{1/2}<(n+1)/2$. Therefore $\sigma(n)/\tau(n)\leqq(n+1)/2$, with equality if and only if n is a prime. This improves the given estimate.

Using this technique, we can also improve the estimate of E1625 [1963, 891, and 1964, 683–4], namely, that $\sigma(n)/\tau(n)\geqq n^{1/2}$. Setting $f'(d)=0$, we see that $f(d)\geqq n^{1/2}$, with equality if and only if $d^2=n$. Therefore $\sigma(n)\geqq n+1+(\tau(n)-2)n^{1/2}$, and so

$$\sigma(n)/\tau(n) \geqq (n^{1/2}-1)^2/\tau(n)+n^{1/2},$$

with equality if and only if n is a prime or the square of a prime. Since $\tau(n)\leqq n$, we see further that

$$\sigma(n)/\tau(n) \geqq n^{1/2}+(1-n^{-1/2}),$$

with equality if and only if $n=2$.

III. *Solution by Leonard Carlitz, Duke University.* We have

$$\sigma(n)/\{n\tau(n)\} = \prod_{p|n} (1 + 1/p + \cdots + 1/p^e)/(e + 1)$$

for $n = \prod_{p|n} p^e$. Since $(1+1/p+ \cdots +1/p^e) \leqq e+1$ and $n>1$, it follows that

$$\sigma(n)/\{n\tau(n)\} \leqq (1 + 1/2 + \cdots + 1/2^e)/(e + 1) \leqq \begin{cases} 3/4 & (e = 1), \\ 2/3 & (e > 1). \end{cases}$$

Hence, $\sigma(n) \leqq 2n\tau(n)/3$ for $n>2$, and of course $\sigma(2) = 3 \cdot 2\tau(2)/4$.

Also solved by J. C. Abad, A. N. Aheart, J. T. Anderson, J. W. Baldwin, Jack Barone, Robert Bernstein, M. A. Bershad, Walter Bluger, W. R. Boland, W. H. Bonney, Dale Burnett, J. A. Burslem, F. A. Butter, Jr., John Christopher, M. R. Chowdhury (Germany), D. I. A. Cohen, G. C. Dodds, R. B. Eggleton (Australia), J. D. Featherstone, N. J. Fine, G. Ford, P. K. Garlick, A. A. Gioia, K. R. Goodearl, Jerry Goodman, H. S. Hahn, D. M. Hancasky, B. A Hausmann, M. H. Hayamizu, Agatha Himmelfarb, Robert Hursey, Geoffrey Kandall, Agnis Kaugars, M. S. Klamkin, Kenneth Kramer, C. D. La Budde, C. C. Lindner, Robert Maas, C. R. MacCluer, E. L. Magnuson, D. C. B. Marsh, C. J. Moreno, J. B. Muskat, M. G. Murdeshwar, C. B. A. Peck, Harsh Pittie, G. B. Purdy (England), Simeon Reich (Israel), G. V. Rowell, M. S. R. K. Sastry, P. A. Scheinok, Robin Sibson (England), Barry Simon, R. Sivaramakrishnan (India), Sidney Spital, E. E. Strock, J. J. Tattersall, Guy Torchinelli, A. M. Vaidya, C. Van de Vyle (Belgium), Emanuel Vegh, Lenard Weinstein, B. A. Welsh, K. L. Yocom, and the proposer.

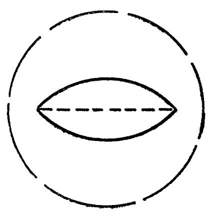
Many solvers deduced the inequality $\sigma(n)/\tau(n) \leqq (3/4)^d n$ where d is the number of distinct divisors of n.

Decomposition of 3-Space by Closed Half-Planes

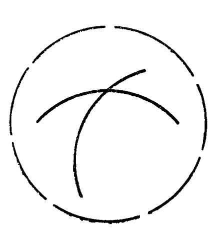
E 1750 [1965, 75]. *Proposed by L. J. Bakanowsky, N. Greenleaf, and M. Walter, Harvard University*

If any two rays, emanating from the origin, are removed from the plane, the plane is cut into two components. What analogous statement(s) can be made with respect to the removal from three dimensional space of closed half-planes whose boundary lines pass through the origin?

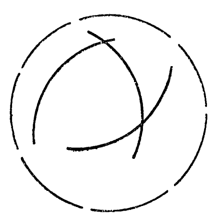
Solution by Michael Goldberg, Washington, D.C.



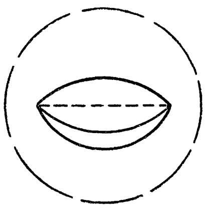
Two half-planes with edge in common; two regions.



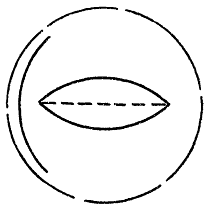
Two half-planes without edges in common; one region.



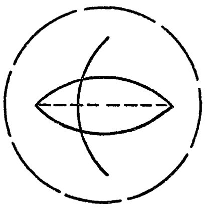
Three half-planes with no edge in common and not cutting each other; one region.



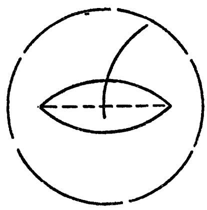
Three half-planes with edge in common; three regions.



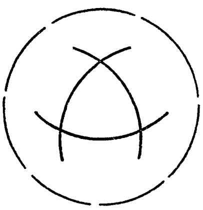
Two half-planes with edge in common, third half-plane entirely in larger region; two regions.



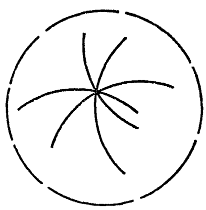
Two half-planes with edge in common, third half-plane dividing the smaller region; three regions.



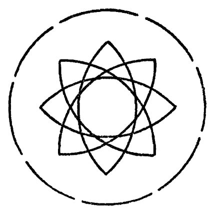
Two half-planes with edge in common, third half-plane in both regions; two regions.



Three half-planes with no edge in common, but cutting each other; two regions.



n half-planes with common intersection; one region.



k pairs of half-planes, each with common edge; $2(k^2-k+1)$ regions.

Also solved by J. W. Baldwin, A. G. Borden, Jr., D. I. A. Cohen, H. S. Hahn, D. M. Hancasky, C. D. LaBudde, Charles McCracken, D. C. B. Marsh, Robin Sibson (England), and the proposer.

$(\text{adj } A)^{\text{tr}} = (\det A)A^{\text{tr}}$

E 1751 [1965, 75]. *Proposed by R. E. Mikhel, Tri-State College, Angola, Indiana*

Let A be an involutory matrix ($A^{-1}=A$). Prove that if every element of A is replaced by its cofactor, then the resulting matrix is an involutory matrix.

Solution by L. A. Gavin, University of Kansas, Lawrence, Kansas. The matrix obtained from A as in the statement of the problem is of course $B=(\text{adj } A)^{\text{tr}}$, and we prove the generalization: If $A^n=I$ for some integer n, then $B^n=I$.

Proof: The conclusion is clear for the case $n=0$. In the remaining cases, A is nonsingular and, so adj $A=(\det A)A^{-1}=(\det A)A^{n-1}$. But then

$$
\begin{aligned}
B^n &= BB^{n-1} = (\det A)(A^{n-1})^{\mathrm{tr}}((\operatorname{adj} A)^{\mathrm{tr}})^{n-1} \\
&= (\det A)((\operatorname{adj} A)^{n-1}A^{n-1})^{\mathrm{tr}} = (\det A)(((\operatorname{adj} A)A)^{n-1})^{\mathrm{tr}} \\
&= (\det A)(((\det A)I)^{n-1})^{\mathrm{tr}} = (\det A)^n I,
\end{aligned}
$$

where we have used the fact that A commutes with its adjoint. The proof is completed by observing that $1=\det I=\det A^n=(\det A)^n$.

Also solved by J. C. Abad, R. G. Albert, D. P. Ambrose (Basutoland), W. H. Bailey, W. T. Bailey, W. C. Barnes, John Beidler, P. M. Berry, M. A. Bershad, Marjorie Bicknell, D. A. Blaeuer, C. F. Blakemore, W. R. Boland, W. H. Bonney, A. G. Borden, Jr., G. R. Bowman, J. F. Burke, J. A. Burslem, David Carlson, Daniel Cohen, R. J. Cormier, W. G. Dotson, Jr., R. B. Eggleton (Australia), E. W. Ewing, J. D. Featherstone, Walter Feibes, N. J. Fine, Philip Fung, P. K. Garlick, L. A. Gavin, D. P. Geller, A. A. Gioia, Robert Gouw, Cornelius Groenewoud, J. D. Haggard and D. W. Height (jointly), H. S. Hahn, C. A. Hall, D. M. Hancasky, Eldon Hansen, J. Z. Hearon, J. C. Hickman, Andrew Jarosak, Agnis Kaugars, S. C. King, M. S. Klamkin, C. D. La Budde, H. E. Lahmann (Germany), E. S. Langford, Rene Laumen (Belgium), T. J. Lee, Steve Ligh, C. R. MacCluer, E. W. Marchand, D. C. B. Marsh, Charles McCracken, C. F. McLaren, Sister Irene Morvan, D. L. Muench, M. G. Murdeshwar, G. L. Musser, J. C. Nichols, Dave Nixon, J. M. O'Neil, J. M. Perry, Harsh Pittie, J. R. Porter, V. K. Rohatgi, Bernard Rosner, W. M. Sanders, Ronald Satterwhite, P. A. Scheinok, Robin Sibson (England), Barry Simon, Indranand Sinha, R. Sivaramakrishnan (India), G. P. Speck, Sidney Spital, R. L. Syverson, E. J. Taft, R. P. Tapscott, Mrs. Elaine Tatham, Rory Thompson, D. P. Tihansky, B. R. Toskey, Elias Toubassi, C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), Emanuel Vegh, Lenard Weinstein, R. A. Wiesen, C. R. Williams, Theodore Zerger, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. *Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before July 31, 1966.*

5350. *Proposed by S. A. Naimpally, Iowa State University*

Menger (Math. Ann. 100 (1928), 75–163) defined a convex metric space (X, d) to be one for which, corresponding to each pair of distinct points $p, q \in X$, there exists a point $r \in X$ different from p, q such that $d(p, q)=d(p, r)+d(r, q)$, and showed that if X is complete then p, q are joined by a metric segment $\overline{pq}$, i.e. by a subset of X which is isometric to a segment of the real line of length $d(p, q)$. Further, if this metric segment is unique we call X a *strongly convex* space.

If (X, d) is a strongly convex complete metric space with a bounded metric d, and F is the function space of all continuous self-mappings of X with metric ρ, where for $f, g \in F$, $\rho(f, g)=\sup_{p\in X} d(f(p), g(p))$, then show that F is convex. Is F convex when X is merely convex?

5351. *Proposed by R. G. Buschman, State University of New York at Buffalo*

Louis Brand (this MONTHLY, 71 (1964) 719–728) discusses a convolution ring S for sequences of complex numbers and its extension into a division alge-

bra. It is not difficult to see that for a sequence $A \in S$ with $a_0 = 0$, the equation $X^2 = A$ has no solution.

(i) Show that $X^2 = A$ has a solution in S if and only if the index of the first nonzero element of the sequence $A \neq 0$ is even.

(ii) Hence, for the Fibonacci sequence, $Y^2 = \{f_n\}$ has no solution, but $X^2 = \{f_{n+1}\}$ has a solution. Extract the square roots of this "left-shifted" Fibonacci sequence.

(iii) Indeed, show that $\{i^{-n} P_n(i/2)\}^2 = \{f_{n+1}\}$, where the P_n are the Legendre polynomials.

5352. *Proposed by Kwangil Koh, North Carolina State University*

A ring is called *duo* if every one-sided ideal of the ring is a two-sided ideal. If R is a ring let R^Δ be the singular ideal of R. Prove that for a duo ring R, the following conditions are equivalent:

(i) $R^\Delta = (0)$.

(ii) R is isomorphic to a subdirect sum of Ore domains. (A ring R is said to be an Ore domain if it is an integral domain and any pair of nonzero right ideals and left ideals has a nonzero intersection.)

(iii) R is isomorphic to a subring of a direct sum of division rings.

5353. *Proposed by Kwangil Koh, North Carolina State University*

A subdirectly irreducible duo ring which is semi-simple is a division ring.

5354. *Proposed by B. R. Toskey, Seattle University*

An exercise in Fuchs, *Abelian Groups*, (ex. 26, p. 36) is to prove that if G is an R-module, where R is a (commutative) domain of integrity, then the elements of nonzero order form a submodule. Is the statement true if R is a noncommutative integral domain?

5355. *Proposed by P. M. Weichsel, University of Illinois*

Let G be a group. Show that u is an element of the commutator subgroup of G if and only if there exist elements $g_1, \cdots, g_k$ in G such that $u = g_1 g_2 \cdots g_k$ and $1 = g_k g_{k-1} \cdots g_1$.

5356. *Proposed by K. E. Whipple, Auburn University, Alabama*

Suppose that F is a subfield of the real numbers and P is a subset of F having the following properties:

(1) P is closed with respect to addition and multiplication.

(2) If x belongs to F, then only one of the following is true, (i) $x = 0$, (ii) $x \in P$, (iii) $-x \in P$.

Must P be the set of positive numbers in F?

5357. *Proposed by David Rearick, University of Colorado*

If n is a positive integer, let $f(n)$ denote the number of ways of expressing n

as a product of two relatively prime positive integers, disregarding order. Let $\nu(n)=(-1)^{r(n)}$ where $r(n)$ is the number of distinct prime divisors of n. Show that

$$\sum_{n=1}^{\infty}\frac{\nu(n)f^3(n)}{n^2}=\frac{7}{8}.$$

5358. *Proposed by H. F. Mattson and E. F. Assmus, Jr., Sylvania Electronic Systems, Waltham, Mass.*

Let A be an $m\times n$ matrix over the integers with $2<m+1<n$. Prove that if p is a prime such that $p<\max(m, n-m)$ then p divides one of the $m\times m$ determinants in A.

5359. *Proposed by Jean-Pierre Serre, Paris, France*

Let A_n be the projective n-space ($n\geqq 2$) over some field K. A finite subset X of A_n is called a Sylvester-Gallai (S-G) configuration if it verifies the following condition:

(*) If P, $Q\in X$, with $P\neq Q$, the line joining P and Q contains at least one more point of X. (Equivalently: no line intersects X in exactly two points.)

An S.-G. configuration is called linear (planar) if it is contained in a line (plane). If K is the field of real numbers, it is known that any S-G configuration is linear. Over the field of complex numbers there are well-known examples of nonlinear S-G configurations (e.g., the nine inflection points of a nonsingular cubic).

Is there a nonplanar Sylvester-Gallai configuration over the field of complex numbers?

SOLUTIONS OF ADVANCED PROBLEMS

3133 [1965, 915]. *Erratum in Note by Albert Wilansky, Lehigh University.* Bottom line, for "identity" read "idempotent."

Ares Bounded by Mutually Tangent Ellipses

4744 [1957, 437]. *Proposed by M. S. Klamkin, University of Minnesota*

Three congruent ellipses are mutually tangent. Determine the maximum and the minimum of the area bounded by the three ellipses.

Solution by Michael Goldberg, Washington, D. C. The extremal positions can be found as equilibrium states of the following hydromechanical analogy. The minimum enclosed area is obtained by uniform external pressure on the ellipses while the maximum enclosed area is produced by a pressure within the enclosed area. In either case, the resultant force due to the pressure on an ellipse is directed along the perpendicular bisector of the chord joining the points of contact. This force can be resisted only by normal forces at the points of contact. To obtain

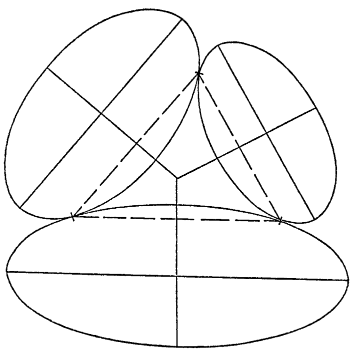
equilibrium, the ellipse will turn until these forces are concurrent. But this can occur only when the chord is parallel to an axis of the ellipse, and this applies to each of the ellipses. Hence, the maximum enclosed area is attained when the arcs of least curvature bound the area: that is, when the chords are parallel to the major axes and form an equilateral triangle. The least area is attained when the chords are parallel to the minor axes and form an equilateral triangle.

The enclosed areas can be derived as follows. Begin with a circle of radius a and two tangents which make an included angle of $2A$. Then the area between the circle and the tangents is $a^2(\cot A+A-\pi/2)$. If the figure is rotated about the bisector of the angle $2A$ so that the circle projects into an ellipse of semiminor axis b, then the included area becomes $ab(\cot A+A-\pi/2)$, but the angle between the tangents becomes $2B$, where $a\tan B=b\tan A$. Thus the minimum area for the given problem is obtained when $B=\pi/3$, $\tan B=\sqrt{3}$, $\tan A=3^{1/2}a/b$, and the total enclosed area is $3ab(b/a\sqrt{3}-\tan^{-1}(b/a\sqrt{3}))$.

If the figure is rotated about the normal to the bisector of angle $2A$ so that the circle projects into an ellipse of semiminor axis b, then the new included area again becomes $ab(\cot A+A-\pi/2)$ but the new angle between the tangents becomes $2B$, where $b\tan B=a\tan A$. For the maximum area of the given problem, $B=\pi/3$, $\tan B=\sqrt{3}$, $\tan A=b\sqrt{3}/a$, and the total enclosed area is $3ab(a/b\sqrt{3}-\tan^{-1}(a/b\sqrt{3}))$.

The foregoing procedures can be used when the three ellipses are not congruent, even when other curves are used. The following theorem may be stated: *A necessary condition for an extremum for the area enclosed by three (or more) mutually tangent curves is the equilibrium condition that the normals at the points of contact of each curve make equal angles with the respective chords joining these points of contact.*

The adjoining figure is an example of the maximum area enclosed by three unequal ellipses.



Also solved by A. R. Hyde.

Derivatives of Functions Continuous on the Rationals

5181 [1964, 325; 1965, 327]. *Proposed by E. J. Burr, University of New England, Australia.*

Let $f(x)=0$ for x irrational, $f(0)=\epsilon_1>0$, and $f(m/n)=\epsilon_n>0$ for m, n coprime integers with $n>0$. Find, or disprove the existence of, sequences $\{\epsilon_n\}$ with $\lim \epsilon_n=0$ such that (a) $f'(x)$ exists nowhere, (b) $f'(x)$ exists for some x, (c) $f'(x)$ exists for all irrational x.

Note by M. D. Mavinkurve. In the original solution [1965, 327] the second sentence of part (c) should read as follows: For by a theorem due to M. K. Fort, the points of continuity of the derivative of a real function (on any interval) which itself is discontinuous on a dense subset, form a set of the first category, and the set of irrationals is not of the first category.

A Filter on a Set of Monotonic Sequences

5234 [1964, 923]. *Proposed by George Bergman, Harvard University*

On the set of sequences $b=(b_0, b_1, \cdots)$, $1\geqq b_i\geqq b_{i+1}\geqq 0$, let F be the filter having the following subbasis: for every convergent series of real numbers $\sum_{i=0}^{\infty}a_i$ and every positive ϵ, the set $\{b: |\sum a_ib_i-\sum a_i|<\epsilon\}\in F$. Find a simple expression for F. Prove therefrom that if $\sum a_i$ converges, then $\lim \sum_{x\to 1}a_ix^i=\sum a_i$.

Solution by the proposer. F may be described as the filter of termwise convergence to the sequence $(1, 1, 1, \cdots)$. This is seen from the fact that a basis for F is the set of sequences $\{b=(b_1, b_2, \cdots)\}$ described as follows: For real $\delta>0$, $B\equiv\{b: b_n>1-\delta$ for every positive integer $n\}$. (In this description we retain the condition $1\geqq b_i\geqq b_{i+1}\geqq 0$.)

Proof. These sets form a filter basis, for the intersection of every finite set of them contains a set of the same sort, namely the one obtained by taking the minimum δ. Each of them is contained in F—is, in fact, the member of our subbasis gotten by taking $a_i=\delta_i^n$, $\epsilon=\delta$. To show the equivalence of the filters generated, it will suffice to show that any member of F contains one of these basis elements, i.e., that given a convergent series $\sum a_i$ and an $\epsilon>0$, there exists a $\delta>0$ such that for $b\in B$, $b_n>1-\delta$ implies $|\sum a_ib_i-\sum a_i|<\epsilon$.

Choose $n\geqq 1$ such that for all $k\geqq n$, $|\sum_{i=k+1}^{\infty}a_i|<\epsilon/3$, and let $\delta=\epsilon/3n\ \max_{0\leqq k}|\sum_{i=0}^{k}a_i|$. Now

$$(\inf b_i)+(b_0-b_1)+(b_1-b_2)+\cdots+(b_k-b_{k+1})+\cdots$$

is an absolutely convergent series (with sum b_0). Hence the series

$$(\inf b_i)\sum_{i=0}^{\infty}a_i+(b_0-b_1)a_0+\cdots+(b_k-b_{k+1})\sum_{i=0}^{k}a_i+\cdots$$

converges, and its sum $=\sum a_ib_i$, which is within ϵ of $\sum a_i$.

For $0 \leqq x \leqq 1$, clearly $(x^0, x^1, x^2, \cdots) \in B$. Then, it suffices to observe that as $x \to 1$, this sequence converges termwise to $(1, 1, \cdots)$.

Addition of Plane Convex Figures

5253 [1965, 84]. *Proposed by Joseph Hammer, University of Sydney, Australia*

"Addition" of plane convex figures ϕ_1 and ϕ_2 is the following operation: Take any point R_1 in ϕ_1, and any point R_2 in ϕ_2, with an arbitrary point in the plane taken as origin O. The point R which is defined as the vector-sum

$$\overrightarrow{OR} = \overrightarrow{OR_1} + \overrightarrow{OR_2}$$

is considered as the sum of the points R_1 and R_2. If for every possible R_1 in ϕ_1 and R_2 in ϕ_2 we construct R as the sum of R_1 and R_2, then the locus of R is called the "sum" of ϕ_1 and ϕ_2.

Let us denote the boundaries of ϕ_1, ϕ_2 and $\phi = \phi_1 + \phi_2$ by K_1, K_2 and K respectively. Prove that the area enclosed by K is greater than the sum of the areas enclosed by K_1 and K_2. Has this area an upper bound?

I. *Solution by P. R. Scott, Victoria University, Wellington, New Zealand.* For any two bounded two dimensional figures ϕ_1 and ϕ_2 in the plane having boundary curves K_1, K_2 respectively, the figure $\phi_1 + \phi_2$ can be obtained in the following way. Let ϕ_1 have a fixed position in the plane, and let A_2 be a fixed point of K_2. Denote by $\phi_2^{(P)}$ the figure obtained by translating ϕ_2 so that A_2 coincides with a point P of K_1. Then the set of all figures $\phi_1 \cup \phi_2^{(P)}$, in which P ranges over all points of K_1, forms a figure which equals the sum $\phi_1 + \phi_2$.

Suppose now that K_1 and K_2 are given the same orientation. Let l_2 be a support line (for definition, see Yaglom and Boltyanskii, *Convex Figures*, p. 7) to ϕ_2 through A_2, and let l_1 be a support line to ϕ_1, parallel to l_2 but with opposite orientation. Let A_1 be a contact point of l_1 and K_1. If ϕ_2 is now translated to position $\phi_2^{(A_1)}$ so that A_1 and A_2 coincide, l_1 and l_2 also coincide; hence ϕ_1 and $\phi_2^{(A_1)}$ have only boundary points in common. Since $\phi_1 \cup \phi_2^{(A_1)}$ is contained in the sum figure, the first required result follows immediately.

It is easy to see that this area has an upper bound. If $\phi = \phi_1 + \phi_2$ has boundary K then it is well known that the length of K is the sum of the lengths of K_1 and K_2 (*loc. cit.*, p. 45). The figure of given perimeter L and maximum area is the circle. Hence the area of the sum figure is bounded above by $(L_1 + L_2)^2/4\pi$, where L_i is the length of K_i. This bound is precise when ϕ is a circle.

II. *Solution by Fritz Steinhardt, City College, City University of New York.* The "addition" in question, $\phi = \phi_1 + \phi_2$, was first studied by H. Brunn and H. Minlowski (see, e.g., Bonnesen and Fenchel, *Theorie der konvexen Körper*, 1934, chapters 5, 11). Denote by $\lambda\phi_i$ the plane convex figure obtained from ϕ_i by magnification with center 0 in the ratio $\lambda : 1$. Then the linear series connecting ϕ_1 and ϕ_2 is the family $\{(1-\lambda)\phi_1 + \lambda\phi_2 \mid 0 \leqq \lambda \leqq 1\}$ of convex figures, which for $\lambda = \frac{1}{2}$ contains $\frac{1}{2}(\phi_1 + \phi_2) = \frac{1}{2}\phi$.

Denote the area of $(1-\lambda)\phi_1+\lambda\phi_2$ by $K_{1+\lambda}$ (hence $K_{3/2}=$area of $\frac{1}{2}\phi=\frac{1}{4}K$). The Brunn-Minkowski theorem on linear series (*loc. cit.* p. 88) implies that $\sqrt{K_{1+\lambda}}$ is a convex function of λ, i.e. that

$$\sqrt{K_{1+\lambda}} \geqq (1-\lambda)\sqrt{K_1}+\lambda\sqrt{K_2}.$$

Hence $K_{3/2}\geqq\frac{1}{4}(K_1+K_2+2\sqrt{K_1K_2})$, or $K=4K_{3/2}\geqq K_1+K_2+2\sqrt{K_1K_2}$, whence $K>K_1+K_2$ unless one at least of ϕ_1, ϕ_2 is a point.

The ratio $K/(K_1+K_2)$ has no finite upper bound. In fact, if ϕ_1, ϕ_2 are nonparallel line segments, then $K_1=K_2=0$ while K is the area of the parallelogram ϕ whose sides are congruent to ϕ_1 and ϕ_2.

Similar results and discussion apply to dimension n, where the Brunn-Minkowski inequality asserts the concavity of the function $\sqrt[n]{K_{1+\lambda}}$ of λ.

Also solved by Mordecai Avriel, E. O. Buchman, Michael Goldberg, Charles McCracken, Robin Sibson (England), S. J. Sidney, F. B. Strauss, Benjamin Volk, R. S. Yerkes, and the proposer.

Monotone Multiplicative Number-Theoretic Functions

5254 [1965, 84]. *Proposed by A. M. Gleason, Harvard University*

Let f be a weakly increasing function from the positive integers to the positive real numbers. Suppose that $f(mn)=f(m)f(n)$ whenever m and n are relatively prime integers. Show that there is a nonnegative number α such that $f(n)=n\alpha$ for all n.

Editorial Note. The solution to this problem with related references appears in connection with problem E 1735 (1965, 912).

Complete solutions were submitted by Y. M. ben-David, Robert Breusch, Yu Chang, N. J. Fine, R. M. Krause, Sim Lasher and Louis Gordon, C. L. Vanden Eynden, Benjamin Volke, R. S. Yerkes, and the proposer. Also acknowledged with thanks are communications from Paul Bateman, L. Carlitz, E. S. Langford, Joseph Lehner, Andrzej Makowski (Poland), A. A. Mullin, M. G. Murdeshwar, C. B. A. Peck, Robin Sibson (England), S. J. Sidney, and Al. Somayajulu, all of which provided references to E 1735 or to other results already listed there.

The proposer's solution contains the following comment and questions. If we add only the hypothesis that f take integer values, we are led to the question: If α is such that n^α is an integer for every integer n, must α be an integer? It is reasonable to conjecture that even more is true, for example: If 2^α and 3^α are both integers, then α is an integer. The latter is subsumed under the following conjecture, a proof of which would be a milestone in number theory: *The logarithms of the primes are algebraically independent over the rationals.*

Jacobians of Elementary Symmetric Functions

5256 [1965, 85]. *Proposed by K. D. Taylor, George Washington University*

Show that the transformation $A:(x_1,\cdots,x_n)\to(\sigma_1,\cdots,\sigma_n)$, where the σ's are the elementary symmetric functions in n variables, has the Jacobian equal to the Vandermonde determinant $\pi(x_i-x_j)$, where i runs from 1 to $n-1$ and j runs through all values greater than i.

I. *Solution by D. P. Ambrose, University of Basutoland.* The Jacobian is the $n\times n$ determinant with ith column,

$$(1, \sigma_1 - x_i, \sigma_2 - x_i\sigma_1 + x_i^2, \cdots, \sigma_{n-1} - x_i\sigma_{n-2} + \cdots + (-1)^{n-1}x_i^{n-1}),$$

which is of degree $\frac{1}{2}n(n-1)$ in $x_1, \cdots, x_n$.

Putting $x_i=x_j$, the determinant vanishes, since two columns are then identical. Thus x_i-x_j is a factor for each of the $\binom{n}{2}$ pairs i, j such that $i\neq j$. This provides the required $\frac{1}{2}n(n-1)$ factors, and hence the determinant is $k\Pi_{i<j}(x_i-x_j)$, where k can be obtained by comparing coefficients of $x_1^{n-1}x_2^{n-2}\cdots x_{n-1}$.

In the determinant this coefficient (by considering the non-leading diagonal of the minor of the least element of the first row) is 1. Thus the Jacobian equals the Vandermonde determinant.

II. *Solution by I. Olkin, Stanford University.* First consider the transformation to power sums, i.e.,

$$y_k = \sum_{i=1}^{n} \sigma_i^k, \qquad k = 1, \cdots, n.$$

The determinant of derivatives is immediately evaluated as $n!\ V(\sigma_1, \cdots, \sigma_n)$, where V is the Vandermonde determinant. To transform $y\to x$, where x_r is the rth elementary symmetric function, note that $y_1^k=y_k+A(y_1, \cdots, y_{k-1})+kx_k$, where $A(y_1, \cdots, y_{k-1})$ is a function of $y_1, \cdots, y_{k-1}$. The matrix of derivatives is triangular, with diagonal elements $1, 2, \cdots, n$, and hence the determinant is $(n!)^{-1}$. The product of Jacobians yields the desired result.

Also solved by M. G. Beumer (Netherlands), Robert Breusch, F. H. Butter, Jr., L. Carlitz, N. J. Fine, A. G. Heinicke, D. A. Hejhal, John B. Kelly, Stanton Philipp, V. L. N. Sarma (India), A. Sharma, S. J. Sidney, Sidney Spital, W. F. Trench, Simon Vatriquant (Belgium), and the proposer.

Beumer points out that another formulation of this problem was presented as no. 4009 [1942, 694].

Continuation of Analytic Functions

5257 [1965, 85]. *Proposed by H. S. Shapiro, University of Michigan*

Show that there exists an infinite set N of positive integers, and a sequence of complex numbers z_n, $|z_n|<1$, such that $\lim |z_n|=1$ and, for every $f(z)=\sum_{n=0}^{\infty} a_nz^n$ analytic in $|z|<1$, if (i) $a_n=0$, $n\notin N$, and (ii) $f(z_n)=0$, $n=1, 2, \cdots$, then $f\equiv 0$.

Solution by P. B. Kennedy, University of York, England. If the set N can be arranged as a strictly increasing sequence $\{n_k\}$ satisfying

$$(1) \qquad n_j \mid n_k \quad \text{for all } j, k \text{ with } j < k,$$

(e.g., $n_k=2^k$), then a sequence $\{z_n\}$ with the required property can be constructed as follows. Let $\{r_k\}$ be a sequence of numbers in $(0, 1)$ such that $r_k\to 1$ as $k\to\infty$,

and put $z_n=0$ $(n<n_1)$, and

$$(2) \qquad z_n = r_k \exp\{2\pi i(n - n_k)/n_k\} \qquad (n_k \leqq n < n_{k+1},\ k = 1, 2, \cdots).$$

It is obvious that $|z_n|\to 1$ as $n\to\infty$.

Suppose now that $f(z)=\sum_{k=1}^{\infty} a_k z^{n_k}$ is analytic in $|z|<1$ and satisfies $f(z_n)=0$ $(n=1, 2, \cdots)$. For each $k>1$ put

$$f_k(z) = \sum_{j=1}^{k-1} a_j z^{n_j}, \qquad g_k(z) = f(z) - f_k(z).$$

It is clear from (1) and (2) that $g_k(z_n)=g_k(r_k)$ for $n_k\leqq n<n_{k+1}$. But by (1), $n_{k+1}\geqq 2n_k$, and so by (2) the set $\{z_n\colon n_k\leqq n<n_{k+1}\}$ consists of n_k distinct points exactly. Hence the polynomial f_k, of degree n_{k-1}, takes the value $-g_k(r_k)$ at n_k distinct points. Thus f_k is independent of z, and because k is arbitrary the conclusion $f\equiv 0$ follows easily.

Also solved by S. J. Sidney, Benjamin Volk, and the proposer who raises the additional questions: (1) Does there exist a set N corresponding to a given sequence $\{z_n\}$ with the required property? (2) What kind of a condition must be set for N before it has an associated $\{z_n\}$? In Volk's example, N is the set of composites.

Invertible Elements in a Finite Ring

5258 [1965, 85]. *Proposed by David Singmaster, University of California, Berkeley*

In any ring, a divisor of zero cannot have an inverse. Show that, in a finite ring, the existence of an element which is not a divisor of zero implies the existence of a unit and implies that any noninvertible element is a divisor of zero. (We consider 0 as a divisor of zero.)

Solution by Kenneth Yanosko, University of Chicago. Suppose that a does not divide zero. Since R is finite there exist integers $m<n$ with $a^m=a^n$. Then $a^{m-1}(a-a^{n-m+1})=0$, so that $a=a^{n-m+1}$. For any $b\in R$, $ba=ba^{n-m+1}$ which implies $0=ba-ba^{n-m+1}=(b-ba^{n-m})a$ which implies $b=ba^{n-m}$. Similarly $b=a^{n-m}b$, so that a^{n-m} is a multiplicative identity. Furthermore, a is invertible, so any noninvertible element must be a divisor of zero.

Also solved by D. J. Allen, G. C. Anderson, D. R. Arterburn, D. W. Bailey, K. F. Bailie, G. Baron & W. Imrich (Austria), H. E. Bell, Y. M. ben David, D. E. Berthoff, Robert Bowen, F. P. Callahan, L. Carlitz, David Carlson, V. C. Cateforis, Paul Chabot, F. B. Crippen, J. R. Durbin, Dianne Erickson, E. W. Ewing, N. J. Fine, L. A. Gavin, Michael Gemignani, R. W. Gilmer, Jr., A. A. Gioia, Myron Goldstein, Paul Grabarkewitz, Charles Green & R. J. Lifsey, Jr., T. W. Hargrave, M. E. Harris, A. G. Heinicke, D. A. Hejhal, G. A. Heuer, W. D. Jackson, J. S. Johnson, K. G. Johnson, Geoffrey Kandall, V. H. Keiser & J. P. Burling, Kwangil Koh, A. G. Konheim, Donald La Budde, E. S. Langford, W. G. Leavitt, Steve Ligh, Kaye Lilja, C. C. Lindner, A. E. Livingston, C. R. MacCluer, D. J. Magnuson, D. C. B. Marsh, J. J. Martinez, M. D. Mavinkurve (India), Charles McCracken, T. Mitchell & F. D. Parker & F. R. Olson, R. C. Mullin, M. G. Murdeshwar, Gautam Pandya, H. A. D. Paris (Netherlands), Harsh Pittie, J. R. Porter, C. R. Rao, J. G. Rau, N. A. Razak, Arnold Reinhold, W. R. Ristow, V. K. Rohatgi, Azriel Rosenfeld, Francis Sandomierski, P. S. Schnare, Ross Schneider, Robin Sibson, S. J. Sidney, Indranand Sinha,

W. T. Smythe, Al. Somayajulu, John Stout, F. B. Strauss, E. J. Taft, T. N. Taylor, W. F. Trench, A. M. Vaidya, C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), Emanuel Vegh, Lenard Weinstein, Carol Williams, Randy Woodward & Richard Tangeman, and the proposer.

J. S. Johnson observes that the result remains true even if R is not finite but satisfies the descending chain condition for both left and right ideals.

Positive Definite Quadratic Forms

5259 [1965, 85]. *Proposed by K. Mahler, The Australian National University, Canberra*

Let $f(x_1, \cdots, x_n) = \sum_{h=1}^n \sum_{k=1}^n a_{hk}x_hx_k$ be a positive definite quadratic form. Assume that, as $(x_{11}, \cdots, x_{1n}), \cdots, (x_{n1}, \cdots, x_{nn})$ run over all sets of n points with integral coordinates,

$$f(x_{11}, \cdots, x_{1n}) \cdots f(x_{n1}, \cdots, x_{nn}) \geqq a_{11}a_{22} \cdots a_{nn} \begin{vmatrix} x_{11} & \cdots & x_{1n} \\ \cdot & \cdot & \cdot \\ x_{n1} & \cdots & x_{nn} \end{vmatrix}^2 .$$

Then f is a diagonal form, i.e., $a_{hk}=0$ if $h\neq k$.

I. *Solution by S. J. Sidney, Harvard University.* First, if $x=(0, \cdots, 0, 1, 0, \cdots, 0)$ with kth entry 1, we get $a_{kk}=f(x)>0$. Thus the a_{kk} are positive.

We claim $a_{hk}+a_{kh}=0$ for $h\neq k$. Suppose not, say $a_{12}+a_{21}\neq 0$. Let m be an integer and set $x_{11}=m$, $x_{12}=1$, $x_{kk}=1$ for $k>1$, and $x_{hk}=0$ otherwise. Then the stated inequality reads $[a_{11}m^2+(a_{12}+a_{21})m+a_{22}]a_{22}a_{33} \cdots a_{nn} \geqq a_{11}a_{22} \cdots a_{nn}m^2$, which reduces to $(a_{12}+a_{21})m+a_{22}\geqq 0$. But appropriate choice of m renders this inequality false. Thus $a_{hk}+a_{kh}=0$ for $h\neq k$.

Now, if $h\neq k$, f is not changed if we add α to a_{hk} and subtract it from a_{kh}. Setting $\alpha=-a_{hk}=a_{kh}$, we see that f can be given by a diagonal matrix.

II. *Solution by Marvin Marcus, University of California, Santa Barbara.* Let $A=(a_{hk})$ and let $x_i=(x_{i1}, \cdots, x_{in})$. If $(\,,)$ denotes the standard inner product in the space of real n-tuples then the hypothesis reads

$$\prod_{i=1}^n (Ax_i, x_i) \geqq \prod_{i=1}^n a_{ii} \det((x_i, x_j)), \tag{1}$$

for all integral vectors $x_1, \cdots, x_n$. It is obvious that the inequality (1) also holds for rational vectors and hence, by continuity, for real vectors. Thus let $x_1, \cdots, x_n$ be an orthonormal set of real eigenvectors for A so that (1) becomes

$$\det(A) \geqq \prod_{i=1}^n a_{ii}.$$

The Hadamard determinant theorem then implies that $\det(A)=\prod_{i=1}^n a_{ii}$. But A is positive definite and equality holds in the Hadamard inequality if and only if H is a diagonal matrix.

Also solved by K. R. Goodearl, W. F. Trench, W. C. Waterhouse, and the proposer.

Incomplete Topological Algebras

5260 [1965, 85]. *Proposed by M. Rajagopalan and A. Wilansky, Lehigh University*

A topological algebra is called a Q-algebra if its set of invertible elements is open. Must a normed Q-algebra be complete?

Solution by Duane W. Bailey, Amherst College. No. Let A be the algebra of all continuous functions on $(-\infty, \infty)$ which are constant except on a compact set (which may depend upon the function), with pointwise operations. Let $\|x\| = \sup_{-\infty<t<\infty} |x(t)|$. It is easy to see that $x \in A$ is invertible if and only if x never vanishes, which implies that the set of invertible elements is open. However, A is not complete, for the sequence $\{x_n(t)\}$, where $x_n(t) = e^{-t^2}$, for $t \in [-n, n]$ and $x_n(t) = e^{-n^2}$ for $t \notin [-n, n]$, is a Cauchy sequence which does not converge in A.

Also solved by D. R. Arterburn, C. R. MacCluer, S. J. Sidney, Seth Warner, W. C. Waterhouse, A. Wilansky, and Kenneth Yanosko.

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University
COLLABORATING EDITORS: K. O. MAY, University of California, Berkeley, and E. P. VANCE, Oberlin College

Materials intended for review should be sent directly as follows: Books: R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457. Programmed Materials: K. O. May, University of California, Berkeley, Calif. 94704. Films: E. P. Vance, Oberlin College, Oberlin, Ohio 44074.

The Language of Nature: An Essay in the Philosophy of Science. By David Hawkins. Freeman, San Francisco and London, 372 pp. $7.50.

This is an ambitious book in the philosophy of science. It not only attempts to cover standard topics in the philosophy of science, but to relate the scientific enterprise to other forms of human activity, including those generally embraced by economics and ethics. It is also a strikingly original book in many ways. A pervasive theme is "that the content of positive knowledge reacts upon the ways of thought from which that knowledge evolved, and even, inevitably, upon the philosopher's conception of his problems."

The attempt to embody this theme in particular analyses of scientific concepts leads to what will be regarded by many as confusions. The distinction between what is known and what is the case is sometimes blurred; it would be easy to object that Hawkins confuses the normative and descriptive or the logical and empirical aspects of things. But such objections would be quite beside the

point; this apparent confusion is a part of his method. There is no confusion here, but a thesis which is illustrated (rather than defended as such) in his refusal to draw these sharp distinctions. Although I disagree with this thesis, I find it an illuminating one.

The book is particularly outstanding in its treatment of entropy and information. These are made key concepts not only in the discussion of physics, but also in the discussion of biological concepts, evolution, learning, and even induction. Hawkins also tries more seriously than most to trace the relations between knowledge and behavior: more than most, both because of his particular intellectual commitments and because of his acute moral sense.

Although it is offered as a text in the philosophy of science, students will find it difficult. Hawkins' use of the concepts of entropy and information in a wide variety of contexts, for example, is one of the most interesting and enlightening aspects of his philosophy of science; and yet is one that will make the book hard going for the student who is untrained in mathematics and physics. His style is not murky; indeed, it often has considerable sparkle. But his use of metaphor, for example will make the book hard for students; students are not generally prepared to bring their imaginations with them to work on science or on the philosophy of science. (E.g.: "The pathway beaten through a good analogy becomes, in the end, an element in a formal deductive system.")

The general plan of the book is the following: The first two chapters are on number and geometry; the interpretations offered cannot be classified either as empirical or as logical, but fall somewhere between. The next three chapters are on motion and analysis, measurement, and the laws of motion, and are fairly standard in viewpoint. Chapter Six offers a slightly novel interpretation of probability as "physical parametrization [*sic*] testable in terms of relative frequency." Chapters Seven and Eight work up to the introduction of thermodynamic concepts and the concept of information. Chapter Nine deals with the relation between probability and credibility, and offers a cyclical approach to the justification of induction. Chapters Ten and Eleven are concerned with the application of the concepts of entropy and information to biology and evolution, and to psychology and behavior. Chapter Twelve concerns individual ethics, and Chapter Thirteen, the relation between economic theory and social choice.

All in all this is a fascinating essay, and well worth the effort it takes to penetrate its occasional novelties of expression. Some will find it a useful text as well; others will find it very hard to use.

H. E. KYBURG, JR., Wayne State University

The Axiomatic Method: An Introduction to Mathematical Logic. By A. H. Lightstone. Prentice-Hall, Englewood Cliffs, N. J., 1964. vii+246 pp. $6.50.

This is an excellent text and an excellent introduction to mathematical logic. It takes the student from the propositional calculus through the very latest results in the theory of theories (the approach to completeness is essentially that

of Henkin; Vaught's and Robinson's completeness tests are established in the last few sections of the book). The prose is streamlined and efficient: words are not wasted. There are a great many thought-provoking exercises at the end of almost every one of the short sections of the book; these provide one of the book's best features.

There is only one topic I find treated poorly, and that is quantification theory. The "definitions" of universal and existential quantification on pp. 30 and 33 will not work as they stand (there is a confusion of use and mention), and there is a more serious matter involved in the attempt to treat quantifiers as kinds of sentential connectives. The "proof" of the theorem on p. 31, for example, is simply irrelevant to the theorem. Indeed quantificational arguments seem to be handled in much too casual a manner: "strip off the quantifiers . . . " and treat the argument as sentential.

But by and large the book is first rate; and it is particularly valuable for the care with which it traces the complicated connections and interconnections between logic and abstract algebra as these fields are now developing.

H. E. KYBURG, JR., Wayne State University

Quantification Theory. By J. A. Faris. Monographs in Modern Logic. Routledge and Kegan Paul, London, and Dover, New York, 1964. 147 pp. $2.10.

Although this volume is not up to the level of its predecessor (Faris, *Truth-Functional Logic*), it performs its function of providing an introduction to the techniques of quantification theory (with identity and definite descriptions) clearly and unpretentiously. The system of natural deduction presented is essentially that of Copi's *Symbolic Logic*, except that it contains special rules concerning definite descriptions.

Two novel forms of singular quantification are introduced: dummy quantification, in contexts of the form $(dx)\ (\cdots x \cdots)$ where 'd' is used to refer "to an individual whose identity (and whose real name) is undetermined; and proper name quantification, in contexts of the form $(ax)\ (\cdots x \cdots)$; one would write 'Socrates is mortal' as '(sx) (Mortal x)'. The introduction of these new forms of quantification entails the introduction of two special quantification rules to deal with them. Since the new forms of quantification, and the two rules introduced to govern them, are introduced only for pedagogic reasons, and are quickly eliminated in favor of standard rules (dummy names remain, but the dummy quantifiers go), it may be questioned whether the added complications are worth the price; I suspect not.

One particularly strong point of the book is the careful discussion of (and the provision of fairly mechanical recipes for) the symbolization of general and multiply general statements.

H. E. KYBURG, JR., Wayne State University

A Collection of Problems of Mechanics. By I. V. Meshcherskii. Pergamon Press, New York, 1965. xii+518 pp. \$12.50.

This is a translation of the twenty-sixth Russian edition (1960) of a remarkable collection of problems in statics, dynamics and the theory of vibrations. Problems, 1363 in number, many with appropriate illustrations, vary from the relatively simple to the tantalizingly complex. Numerical data and answers have been changed from the metric to the British system of units to meet the needs of the English-speaking market. Answers are given for all problems.

No textual material accompanies the problems, but occasionally a hint is given to guide the student toward the solution. In addition to problems relating to statics in two and three dimensions, rigid body motion, kinematics of particles, dynamics of a particle and of particle systems, there are some 122 problems illustrating vibrating systems. The student is expected to be familiar with Lagrange's equations, and many problems require this technique.

The book can be recommended as supplementary material for students in many varied courses in mechanics, both at the undergraduate and at the graduate level.

S. W. McCuskey, Case Institute of Technology

Analysis, Volume I. By Einar Hille. Blaisdell, New York, 1964. 640 pp. \$9.50.

This book should be in the hands of every teacher of calculus. Almost every topic is treated in a fresh and vigorous manner, which adds new insights to the most familiar material. The style is lively and stimulating with historical comments of great interest. The theory is handled with rigor but without mortis. There are over 2000 problems, many culled from British, French, Russian and Swedish sources.

One of the more novel features is the introduction of complex numbers and functions of a complex variable with a proof of the fundamental theorem of algebra and an excellent treatment of rational functions.

A second novelty is the treatment of the algebra of different classes of functions (e.g., functions continuous on a closed interval and differentiable functions). The class of piecewise convex functions is carefully studied from the point of view of differentiability.

There is an unusually elegant chapter on curve tracing and appendices on Analytic Geometry and Functions of Bounded Variation and Non-Differentiable Functions.

As an example of the delightful surprises in store for the reader, we find:

"Within the range of the prime number tables

$$li(x) - \pi(x)$$

is positive, but it can be proved that the difference changes its sign infinitely often and that the first change of sign occurs for an $x > 10^9$ but not exceeding

$10^{10^{10^{34}}}$. The latter appears to be the largest number that ever served any definite purpose in mathematics."

Although the book treats many topics customarily reserved for a course in advanced calculus, it explains the more traditional ones with great clarity and there are an adequate number of simple problems. The text is intended for "honors students and future mathematics majors." Even here it needs to be used with considerable discretion. In the hands of a student with sufficient mathematical maturity, it would be a liberal education in mathematics. The book is destined to become a classic quite apart from its possible use as an introductory text.

D. E. RICHMOND, Williams College

Generalized Functions and Direct Operational Methods, Volume 1. By T. P. G. Liverman. Prentice-Hall, Englewood Cliffs, N. J., 1964. xii+338 pp. \$10.50.

This should be a rewarding text for use with a class of mature undergraduate mathematicians or exceptionally well-prepared physicists and engineers. It provides a sound introduction to the theory of generalized functions (g.f.'s) and promises interesting applications in the forthcoming second volume.

The author's principal concern is with the g.f.'s of class $\mathfrak{D}'_+$. Roughly speaking these are (equivalence classes of) sequences $\{f_n\}$ of piecewise continuous functions carried on forward half-lines $[a, \infty)$ such that $\langle f_n, \phi\rangle = \int_{-\infty}^{\infty} f_n(t)\ \phi(t)\ dt$ converges for every $\phi \in \mathfrak{D}_-$, the space of C^∞ functions carried on backward half-lines $(-\infty, b]$. Definitions and first properties of such objects occupy Chapter II. Their deeper properties are developed in Chapters V, VI, while the last two chapters, VII, VIII, treat the Laplace transform, and a class of periodic g.f.'s together with Fourier series. The book is amply sprinkled with stars and double stars indicating material of greater difficulty; in particular, the last four chapters are so marked.

The one aspect of the book that warrants criticism is the large number of pages, about 100, that are devoted to linear differential equations with constant coefficients (d.e.'s). Presumably this was included to round off an "unstarred" course in g.f.'s for students who have little technical knowledge beyond ordinary calculus. But for such students even Chapter II must prove quite difficult; and there exist shorter routes to d.e.'s. Physically oriented students might better spend their time on vector-space theory, a context in which the theory of linear d.e.'s with constant coefficients can be stated in two lines:

$$\vec{x}'(t) = A\vec{x}(t) + \vec{f}(t), \qquad \vec{x}(t_0) = \vec{x}_0 \text{ iff}$$
$$\vec{x}(t) = e^{(t-t_0)A}\left\{\vec{x}_0 + \int_{t_0}^{t} e^{(t_0-u)A}\vec{f}(u)\, du\right\}.$$

The book concludes with an appendix containing short tables of operator inverses and Laplace transforms, and a bibliography of 58 items.

J. G. WENDEL, University of Michigan

Discrete and Continuous Boundary Problems. By F. V. Atkinson. Academic Press, New York and London, 1964. xiv+570 pp. $16.50.

Most of this excellent book is concerned with an elementary and detailed analysis of certain symmetric linear operators which arise from boundary value problems for ordinary differential, difference, and recurrence relations. Some non-self-adjoint problems are also treated. The author's aim is to compare the discrete with the continuous, and then to combine the two in a unified treatment. To quote from the Preface: "We shall pursue our task from three directions. We shall present the theory of certain recurrence relations in the spirit of the theory of boundary problems for differential equations. Second, we shall present the theory of boundary problems for certain ordinary differential equations, emphasizing cases in which the coefficients may be discontinuous, or may have singularities of delta-function type. Finally, we give some account of theories which unify the topics of differential and difference equations, relying mainly on the method of replacement by integral equations."

In keeping with the author's goal of making everything as elementary as possible, he has intentionally avoided the use of functional analysis—in particular, the spectral theory of self-adjoint operators in a Hilbert space. For example, rather than apply the abstract characterization of self-adjoint extensions of symmetric operators in terms of isometries of appropriate deficiency spaces, the author has chosen to classify his self-adjoint problems by using the invariance of appropriate quadratic forms. Such forms occur naturally in the Lagrange identity and its analogs.

Another reason given for not treating this material as a subset of a more abstract theory is that a study of boundary problems can still suggest new directions for research in functional analysis. Indeed, it is in the development of this idea that the book attains its most characteristic and individual flavor. Starting from an analysis of the simplest boundary problem $y'=i\lambda y$, $y(0)=y(1)$, on $0\leqq x\leqq 1$, the author progresses at a leisurely pace through examples of recurrence formulas, discrete and continuous Sturm-Liouville theory, orthogonal polynomials, regular and singular problems, scalar and matrix equations, integral equations involving Stieltjes integrals. At each stage analogies with preceding work are pointed out, and often the new step is motivated by physical problems. The singular problems are treated as limiting cases of regular problems; the Weyl limit-point, limit-circle idea is applied to many situations to obtain eigenfunction expansion results.

The book is almost self-contained. There is a preliminary motivating chapter, followed by twelve chapters and six appendices. A list of books and monographs, and an extensive set of notes for each chapter are also included. In the latter, references to research literature and related problems are given. Finally there is a large collection of exercises, which makes the book valuable as a possible text.

E. A. CODDINGTON, University of California, Los Angeles

Functional Analysis. By Albert Wilansky. Blaisdell, New York, 1964. xvi+291 pp. $10.50.

This book is an excellent introduction to the subject. The author has chosen to present the material with close to minimal hypotheses, and the emphasis is therefore on general properties of linear topological spaces rather than on a detailed analysis of special cases such as Hilbert space. The exposition is very clear, partly because of an extensive use of clarifying comments set off in "proof brackets" from the main arguments. A student beginning in functional analysis may well find this book the most accessible in the field.

The first eight chapters are more elementary than the last six and make use of topology only to the extent of metric spaces. Introductory chapters on Banach and Hilbert spaces are among these. Chapter Nine is a self-contained introduction to topology using nets. Linear topological spaces are defined in Chapter Ten and the next three chapters are devoted to their properties including local convexity, duality, and the closed graph theorem. The concluding chapter contains an introduction to Banach algebras.

There are many examples and applications, including a proof of the Riemann mapping theorem, the existence of a continuous function with divergent Fourier series, and the Wiener theorem on the reciprocal of a function with absolutely convergent Fourier series. Other examples appear in the more than 2000 exercises which are one of the strongest features of the book. Most of these are quite easy, but the range is wide.

There are a number of misprints, none of which seems to be serious.

DUANE W. BAILEY, Amherst College

First Course in Functional Analysis. By Casper Goffman and George Pedrick. Prentice-Hall, Englewood Cliffs, N. J., 1965. xi+282 pp. $12.00.

This book contains an introduction to the study of the topology of metric spaces, Banach spaces including L_p, and Hilbert spaces. It differs from other books at this level in that it also includes locally convex topological vector spaces and Banach algebras. The book does not contain spectral theory of operators nor does it contain the theory of semigroups.

An attractive feature of the book is that each major theorem is followed by one or two illustrations of its use. There are many well-chosen problems. Some problems ask the reader to apply an abstract idea to a concrete situation and others ask him to develop more theory.

The book would be better if the references to the bibliography were more frequent and more specific. The main defect of the book is its high cost.

G. W. HEDSTROM, University of Michigan

Numerical Methods Using FORTRAN. By L. Dale Harris. Merrill, Columbus, Ohio, 1964. xi+244 pp. $6.95.

FORTRAN Programming (II and IV). By L. Dale Harris. Merrill, Columbus, Ohio, 1964. x+146 pp. $3.25.

The first of these two books contains a five chapter elementary description of some common numerical methods interwoven with a four chapter minimal description of FORTRAN, together with two appendices on FORTRAN, a 43 entry bibliography, and an index. The second book is essentially the FORTRAN portion of the first book.

The organization of the first book is generally good. The numerical methods are illustrated by flow charts and programs, and these serve to clarify concepts in FORTRAN. Theoretical discussion is minimal and error analysis is lacking; in the chapter on systems of linear equations, the existence and treatment of singular systems are not mentioned.

The second book is not sufficiently clear and comprehensive to constitute a manual. Illustrative examples are scarce, particularly in Chapter 2; between pages 16 and 37 there are 29 references to an example on page 10. Incomplete definitions lead to surprises as in the middle of page 37. Insufficient editing has left many awkward and incorrect sentence formations and notational inconsistencies.

These two books contain a (mostly) well organized, hastily edited, minimal discussion of FORTRAN and numerical methods. The good bibliography will help the interested reader pursue many topics in more depth.

J. B. JOHNSTON, General Electric Research Laboratory

Numerical Methods in FORTRAN. By John M. McCormick and Mario G. Salvadori. Prentice-Hall, Englewood Cliffs, N. J., 1964. 324 pp. $10.75.

This book contains an effective combination of numerical methods and FORTRAN. The methods are clearly and concisely derived, and well illustrated, in the first half of the book; the second half contains 53 complete, working programs incorporating the methods and illustrating the features of FORTRAN. The typography and editing are excellent.

Chapter 1 contains a brief, well-illustrated description of FORTRAN through expressions, function names, IF, and DO.

Chapters 2 through 8 constitute a lucid, concise, elementary text on: interpolation and extrapolation; algebraic and transcendental equations; simultaneous linear equations; initial-, characteristic-, and boundary-value problems; and partial differential equations. All methods are derived analytically, many from Taylor series expansions; some error analysis is given, along with references. Each of the 35 sections has approximately two examples, illustrated with excellent tables and graphs; numerous references are made to associated programs in the second half of the book.

Most of the 53 FORTRAN programs are accompanied by detailed flow charts and copious notes on program functioning and on FORTRAN details. One minor flaw: neither the text nor Program 9.7 seems to treat attempted inversion of singular matrices. Although not a FORTRAN manual in the usual sense, this book can yield a good working knowledge of FORTRAN.

J. B. JOHNSTON, General Electric Research Laboratory

Projective and Related Geometries. By Harry Levy, Macmillan, New York, 1964. x + 405 pp. $11.00.

By now, there are three different approaches to group theoretical treatment of geometry. The first one (due to Klein) defines a geometry as the theory of the invariants of a subgroup of the general linear group. A modernized version of this should add: general linear group over a field and, in the case of a topological field, closed subgroup. The second approach, due to Elie Cartan, is the basis of all of modern differential geometry. Even though it started as an elaboration of Klein's approach and a fusion of Klein's and Lie's ideas, Cartan's methods yield essentially new insights, methods, and geometries. Unfortunately, there exists no exposition of his approach on a reasonably elementary level in global (as opposed to differential) geometry, except for an article by H. Freudenthal and H.-G. Steiner (Gruppentheorie und Geometrie, in: Grundzüge der Mathematik, vol. II, chap. 10a. Vandenhoek & Ruprecht, Göttingen, 1960). That article is strongly recommended to the attention of future authors of textbooks on the foundations of geometries. The third approach goes back to a series of papers by Ch. Wiener in the Leipziger Berichte in the early 1890's, and has been elaborated by G. Thomsen (see, e.g., The treatment of elementary geometry by a group calculus. Math. Gazette, 17, 1933, p. 232) and was recently the subject of a beautiful book (F. Bachmann, Die Entwicklung der Geometrie aus dem Spiegelungsbegriff, Springer, Berlin 1959). This approach also has shown itself to be much more fruitful in the generation of new geometries than the Klein approach. It therefore should be understood that even the most modern exposition of the Klein program represents a stage of mathematics that is no longer a center of active research.

This being said, the book contains about everything in theorems *about* geometries one could ask for. (It contains very few theorems *in* geometry.) In addition, it is self-contained as far as linear algebra over abstract fields is concerned, and is reasonably complete in abstract quadratic algebra. Most probably it is the best book on the market for a rigorous common development of linear algebra and analytic geometry and as such is highly recommended for any non-geometer teaching a beginning graduate course in these subjects.

H. W. GUGGENHEIMER, University of Minnesota

Projective Geometry. By H. S. M. Coxeter. Blaisdell, New York, 1964. xiii+163 pp. $5.50.

From the preface it seems that the author wanted to write a book on projective geometry for high schools. In this he certainly failed, but on the way he produced an excellent text for a one semester course in projective geometry on the advanced undergraduate level. Even though projective geometry is basically 19th century mathematics, the treatment given here is modern in spirit because it uses an axiom system that is not categorical. (At a few places the generality is bought by disregarding cases that can appear in some theories only. This is done by saying: We restrict our attention to · · · , so it cannot be called cheating.) The approach is synthetic and the conciseness of the proofs will cause some difficulties to inexperienced instructors and students. But both will be rewarded for their toils by an understanding of linear methods. The choice of the axiom system as well as the shortness of the book limit the choice of topics. In a last chapter the analytical theory of projective spaces will need quite a bit of additional explanation by the instructor. Unfortunately, the bibliography lists English language titles only.

H. W. GUGGENHEIMER, University of Minnesota

Geometry and Analysis of Projective Spaces. By C. E. Springer. Freeman San Francisco and London, 1964. xi+299 pp. $7.50.

This is a "down to earth" treatment of projective geometry, by which I mean that there is no celebration of abstract formalisms and axiomatic gimmicks attempted, but a natural and easily understandable development of problems, ideas, methods and results. Geometry, analysis and algebra are often simultaneously built up, leading to a balanced treatment of classical projective geometry which will certainly be of great appeal to the interested student. Moreover, although—or rather, since—the book is clearly rooted in the classical spirit of 1600–1900, many modern more "abstract" concepts can be introduced and illustrated in class by a skilled teacher using this book as a teaching guide. The book is carefully written, containing the well-known theorems of projective geometry, plus many "bridges" with metric geometry, analysis and algebra. I would like to emphasize that quite a few topics—like invariants, covariants, jacobians and hessians—are considered, topics which belong to the "classical heritage" of mathematics and which are—at least in their original context—too often ignored and neglected in our sheaf-theoretic age. I believe that the book will fulfill its purpose of being a fine textbook for an introductory course in projective geometry.

ALFRED AEPPLI, University of Minnesota

Axiomatic Analysis. By Robert Katz, Heath, Boston, 1964. 352 pp. $8.50.

In this book the author gives a highly individualized account of elementary logic, set theory, and the real number system. Approximately a third of the book is devoted to the first two of these topics and the last two thirds to an axiomatic treatment of the real numbers.

The text is broken down into 65 lessons made up of 12–15 hundred numbered units. The units are generally quite short, varying from a single sentence to a third of a page. The result is that the book reads almost like a programmed text.

The author has taken considerable pains to eliminate from the reader's mind all possible ambiguities in language and mistakes in logic. His definitions are careful and well-illustrated. Common pitfalls are pointed out and alternative usages are often mentioned. In short, the author has done all he can to make the book clear and self contained. There are exercises throughout, all of which are considered by the author to be an integral part of his development.

The book could be used as a text at the high school or freshman-sophomore college level for courses in elementary logic, elementary algebra, or the real number system. Perhaps the best use would be for self study by a bright high school or college freshman student.

The typography is good and the reviewer caught no misprints.

J. B. ROBERTS, Reed College

Introduction to Higher Algebra. By A. Mostowski and M. Stark. Translated by J. Musielak. Macmillan, New York, 1964. 474 pp. $6.50.

This book covers those topics normally found in standard textbooks on the theory of equations, with additional topics in combinatorics and matrices. The text begins with an introductory chapter on sets, functions and induction followed by a chapter on permutations and combinations. An introduction to fields and complex numbers comprises Chapter 3, while Chapter 4 deals with determinants and their expansion. Chapter 5 includes just enough development of vector spaces to enable the authors to define the rank of a matrix as the dimension of the space spanned by its columns. Linear systems are also included in this chapter. Chapers 6, 7 and 8 consider polynomials in one variable. Topics include the arithmetic of polynomials, multiple roots, interpolation formulae, partial fraction decompositions, separation of roots, cubics and quartics, irreducibility criteria and a brief introduction to algebraic and transcendental numbers. The fundamental theorem of algebra is proved with the use of the basic notions of analysis. Symmetric polynomials and elimination theory are covered in Chapters 9 and 10, and the final chapter is concerned with quadratic and hermitian forms. From three to eight exercises, primarily non-computational, are included in most sections. The translation and typography are adequate.

E. A. MAIER, University of Oregon

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Associate Secretary, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor Oscar Zariski, Harvard University, has been awarded the 1965 National Medal of Science.

Professor Harold Bacon, Stanford University, received the Lloyd W. Dinkelspiel Award for Outstanding Service to Undergraduate Education. The Award was in the amount of $3,500.

Professor Truman Koehler, Muhlenberg College, represented the Association at the inauguration of W. D. Lewis as President of Lehigh University on October 10, 1965.

Rev. E. A. Sharp, S.J., Creighton University, represented the Association at the inauguration of L. E. Traywick as President of the Municipal University of Omaha on October 15, 1965.

Arlington State College: Assistant Professor R. L. Tennison, East Central State College, Dr. Larry Heath, University of Kansas, Dr. D. R. Falconer, University of Texas, and Mr. R. G. Dean, Texas Christian University, have been appointed Assistant Professors; Assistant Professors Robert Coke and John Perryman have been promoted to Associate Professors; Mr. Joe Gilbreath and Mr. L. G. Shilling have been promoted to Assistant Professors.

Ashland College: Assistant Professor A. G. Poorman has been promoted to Associate Professor; Mr. R. L. Wendling has been promoted to Assistant Professor.

Bowling Green State University: Assistant Professor D. E. Ryan, Eastern Michigan University, and Dr. Wallace Terwilliger, Washington State University, have been appointed Assistant Professors; Assistant Professors Clifford Long and Ralph Townsend have been promoted to Associate Professors.

California State College, Fullerton: Assistant Professor R. V. Benson, Long Beach State College, has been appointed Associate Professor; Associate Professor R. C. Gilbert has been promoted to Professor.

California State College, Long Beach: Dr. F. A. Butter, Jr., Aerospace Corporation, El Segundo, California, has been appointed Associate Professor; Associate Professor A. H. Smith has been promoted to Professor; Mrs. Jean L. Conroy has been promoted to Assistant Professor.

California State College, San Bernardino: Professor J. E. Hafstrom, University of Minnesota, has been appointed Professor and Chairman of the Mathematics Department; Mr. D. E. McLeod, University of California, Riverside, has been appointed Lecturer.

University of California, Davis: Professor A. J. Macintyre, University of Cincinnati, has been appointed Visiting Professor; Associate Professor R. F. DeMar, Miami University, and Dr. Kurt Kreith, Arms Control and Disarmament Agency, Washington, D. C., have been appointed Associate Professors; Assistant Professor D. O. Cutler, Texas Christian University, has been appointed Assistant Professor.

University of California, Los Angeles: Dr. David Sanchez, University of Chicago, has been appointed Assistant Professor; Associate Professor P. B. Johnson has been promoted to Professor; Professor Angus Taylor has been appointed Vice President for Academic Affairs.

Central Michigan University: Associate Professor E. H. Whitmore, San Francisco State College, has been appointed Professor and Head of the Mathematics Department; Associate Professor Julia Adkins has been promoted to Professor.

Colorado College: Mrs. Sandra N. Hilt, University of North Carolina, has been appointed Assistant Professor; Associate Professor George Simmons has been promoted to Professor; Professor W. Y. Gateley has been appointed Chairman of the Mathematics Department.

Duke University: Dr. W. E. Sewell, Associate Director of Special Research in Numerical Analysis, has been appointed Adjunct Professor; Dr. Jacob Burlak, University of Glasgow, has been appointed Associate Professor; Dr. R. E. Hodel, Duke University, has been appointed Assistant Professor; Associate Professor S. L. Warner has been promoted to Professor; Professor A. O. Hickson retired on September 1, 1965, with the title of Professor Emeritus.

Georgia State College: Assistant Professors William Leonard, Susquehanna University, and Kenneth Whipple, Auburn University, have been appointed Assistant Professors.

Iowa State University: Dr. J. A. Dyer, Southern Methodist University, and Dr. Peter Colwell, University of Minnesota, have been appointed Assistant Professors.

Louisiana Polytechnic Institute: Professor W. E. Koss, Mississippi State University, has been appointed Professor; Associate Professors J. B. Garner and J. D. Gilbert have been promoted to Professors.

Marquette University: Assistant Professors Miriam E. Connellan, Willard Lawrence, and J. E. Simpson have been promoted to Associate Professors.

University of Montana: Professor R. E. Johnson, University of Rochester, has been appointed Professor; Assistant Professor J. A. Peterson has been promoted to Associate Professor.

New Mexico State University: Associate Professor D. G. Johnson, Pennsylvania State University, has been appointed Visiting Associate Professor; Assistant Professors G. S. Rogers, University of Arizona, and Louis Solomon, Rockefeller University, have been appointed Associate Professors; Associate Professor J. M. Adams, Texas Western College, has been appoined Part-time Associate Professor and Director of the Computing Center; Drs. Joaquin Loustaunau, University of Illinois, Charles Swartz, University of Arizona, and Donald Stevens, New York University, have been appointed Assistant Professors; Associate Professor R. J. Wisner has been promoted to Professor; Assistant Professors E. D. Gaughan and J. D. Thomas have been promoted to Associate Professors; Professor Walter Heinzman retired on June 7, 1965 with the title of Professor Emeritus.

New York University: Professor Ralph Niemann, Colorado State University, has been appointed Visiting Professor; Dr. Bertrand Levy, Office of Naval Research, Washington, D. C., has been appointed Associate Professor; Mrs. Phyllis Strauss, Columbia University, and Dr. Hilbert Levitz, Williams College, have been appointed Assistant Professors; Associate Professors Jerome Berkowitz and Albert Blank have been promoted to Professors; Assistant Professor Anneli Lax has been promoted to Associate Professor.

Newark College of Engineering: Associate Professors A. E. Foster, Carl Konove, and Charles Koren have been promoted to Professors.

North Carolina State University at Raleigh: Associate Professor P. A. Nickel, Montana State College, has been appointed Associate Professor; Dr. R. E. Chandler, Duke University, has been appointed Assistant Professor; Assistant Professor C. H. Little, Jr., has been promoted to Associate Professor.

North Texas State College: Assistant Professor M. R. Hagan, S. F. Austin State College, has been appointed Assistant Professor; Assistant Professor W. D. L. Appling has been promoted to Associate Professor.

Northeast Missouri State Teachers College: Mr. R. A. Knight, University of Nebraska, has been appointed Assistant Professor; Mr. Samuel Lesseig has been promoted to Assistant Professor.

Northwestern University: Assistant Professor H. F. Kreimer, Florida State University, has been appointed Visiting Associate Professor; Assistant Professor Marvin Shinbrot, University of California, Berkeley, has been appointed Associate Professor.

College of Notre Dame of Maryland: Sister Marie Augustine Dowling has been promoted to Assistant Professor and appointed Chairman of the Mathematics Department; Sister Mary Cordia Karl retired September 1965 with the title of Professor Emeritus.

Ohio University: Dr. C. B. Mehr, International Business Machines, San Jose, California, and Dr. K. E. Eldridge, University of Colorado, have been appointed Assistant Professors; Associate Professor R. K. Butner has been promoted to Professor.

Old Dominion College: Assistant Professor K. J. Davis, University of Tennessee, and Dr. H. W. Baeumler, Ohio State University, have been appointed Associate Professors.

Pennsylvania State University: Professor P. C. Hammer, University of Wisconsin, Madison, has been appointed Professor and Chairman of the Department of Computer Science; Dr. Marllyn Boswell, University of Missouri, and Dr. H. L. Shapiro, Purdue University, have been appointed Assistant Professors; Professor Walter Gordon retired on October 1, 1965 with the title of Professor Emeritus.

Polytechnic Institute of Brooklyn: Associate Professor George Bachman has been promoted to Professor; Dr. Stanley Preiser, United Nuclear Corporation, White Plains, New York, has been appointed Assistant Professor; Professor L. A. MacColl retired on September 1, 1965.

Portland State College: Assistant Professors F. S. Cater, University of Oregon, and V. C. Williams, Reed College, have been appointed Assistant Professors.

Purdue University: Professor G. E. Baxter, University of California, San Diego, has been appointed Professor; Dr. F. R. DeMeyer, University of Oregon, has been appointed Assistant Professor; Associate Professor M. L. Keedy has been promoted to Professor; Assistant Professor M. W. DeJonge has been promoted to Associate Professor.

St. Mary's University: Col. John Armitage and Mr. F. J. Carter have been promoted to Assistant Professors; Brother Frank Gutting has been appointed Chairman of the Mathematics Department.

Seton Hall University: Professor Volodymyr Bohun-Chudyniv, Morgan State College, has been appointed Professor; Mrs. Marcelle Friedman, New York University, has been appointed Assistant Professor.

Southeastern Louisiana College: Mr. K. G. Holder, Chrysler Space Division, New Orleans, Louisiana, has been appointed Assistant Professor; Mr. J. R. Weaver has been promoted to Assistant Professor.

Southeastern Massachusetts Technological Institute: Assistant Professor Fred Wolock, Boston College, has been appointed Associate Professor; Dr. A. J. Chandy, Boston University, has been appointed Assistant Professor.

State University of New York at Buffalo: Professor Rafael Artzy, Rutgers, The State University, has been appointed Professor; Dr. Anthony Ralston, Stevens Institute of Technology, has been appointed Professor; Dr. Federico Gaeta, University of Madrid, has been appointed Professor; Dr. E. W. Wallace, University of Leeds, England, has been appointed Associate Professor; Dr. Ubiratan D'Ambrosio, Brown University, has been appointed Assistant Professor; Dr. Theodore Mitchell has been promoted to Associate Professor; Dr. Gerald Itzkowitz has been promoted to Assistant Professor.

University of Tennessee: Professor Herta T. Freitag, Hollins College, has been appointed Visiting Professor; Associate Professor R. M. McLeod, American University of Beirut, has been appointed Associate Professor; Assistant Professors R. J. Bean, University of Wisconsin, and D. W. Lick, University of Redlands, have been appointed Assistant Professors; Assistant Professors D. R. Brown, D. R. Hayes and H. L. Lee have been promoted to Associate Professors.

Trinity College, Connecticut: Dr. G. A. Anderson has been promoted to Assistant Professor; Associate Professor E. F. Whittlesey has been promoted to Professor.

Ursinus College: Assistant Professor Blanche Schultz has been promoted to Associate Professor; Professor F. L. Dennis has been appointed Head of the Department of Mathematics.

University of Washington: Visiting Assistant Professor J. V. Ryff has been appointed Assistant Professor; Dr. David Topping, University of Chicago, has been appointed Assistant Professor; Associate Professor Robert Blumenthal has been promoted to Professor; Assistant Professor Jack Segal has been promoted to Associate Professor.

Western Michigan University: Professor Emeritus E. H. Rothe, University of Michigan, has been appointed Professor; Assistant Professor D. R. Lick, New Mexico State University, has been appointed Associate Professor; Dr. J. D. Tarwater, University of New Mexico, has been appointed Assistant Professor; Assistant Professors J. W. Petro and J. E. Vollmer have been promoted to Associate Professors; Professor C. H. Butler retired with the title of Professor Emeritus.

University of Wisconsin, Madison: Visiting Professor I. J. Schoenberg has been appointed Professor; Dr. R. A. Brualdi, Syracuse University, Dr. J. E. Hall, University of Wisconsin, Madison, Dr. J. J. Roseman, New York University, Dr. D. F. Shea, Jr., Syracuse University, and Dr. M. C. Thornton, University of Illinois, have been appointed Assistant Professors; Associate Professors W. S. Bicknell and S. V. Parter have been promoted to Professors; Assistant Professors Richard Askey, Michael Bleicher and D. W. Crowe have been promoted to Associate Professors.

U. S. Army, Mathematics Research Center, University of Wisconsin: Professor P. R. Masani, Indiana University, has been appointed Visiting Professor; Dr. H. F. Karreman has been promoted to Professor; Professor L. B. Rall has been promoted to Assistant Director of the Mathematics Research Center.

Yale University: Professor A. H. Stone, University of Rochester, has been appointed Visiting Professor; Associate Professor G. B. Seligman has been promoted to Professor; Assistant Professors F. J. Hahn and R. H. Szczarba have been promoted to Associate Professors; Dr. E. L. Stout has been promoted to Assistant Professor.

Professor W. M. Boothby, Washington University, is on sabbatical leave of absence for 1965–66 and will spend the year at the University of Geneva.

Dr. R. K. Brown, U. S. Army Electronics Command, Fort Monmouth, New Jersey, has been appointed Professor and Chairman of the Mathematics Department at Kent State University.

Dr. J. B. Burling, University of Colorado, has been appointed Associate Professor at State University College at Oswego.

Professor Elizabeth Carlson, University of Minnesota, has been appointed Visiting Professor at Macalester College.

Dr. P. J. Chase, California Institute of Technology, has been appointed Assistant Professor at College of Wooster.

Mr. Arthur Eade, Hamden High School, Hamden, Connecticut, has been appointed Assistant Professor at Quinnipiac College.

Mr. R. W. Feldmann, State University of New York at Buffalo, has been appointed Assistant Professor at Lycoming College.

Assistant Professor John Greever, Harvey Mudd College, has been promoted to Associate Professor.

Assistant Professor J. T. Howson, Jr., Rose Polytechnic Institute, has been appointed Assistant Professor at the University of Cincinnati.

Professor W. E. Jenner, University of North Carolina, has been appointed Chairman of the Mathematics Department.

Assistant Professor T. F. Kimes, Austin College, has been promoted to Associate Professor.

Assistant Professor R. S. Kleber, St. Olaf College, has been promoted to Associate Professor.

Mr. J. E. Lightner, Western Maryland College, has been promoted to Assistant Professor.

Dr. A. V. Martin, University of Ibadan, Nigeria, has been appointed Professor at Union College.

Mr. Morton Mecklosky, Suffolk County Community College, has been promoted to Assistant Professor.

Mr. R. T. Morgan, State University College at Fredonia, has been appointed Assistant Professor at Montclair State College.

Miss Edith Moss, Suffolk University, has been promoted to Assistant Professor.

Professor Paul Patterson, Appalachian State Teachers College, has been appointed Chairman of the Mathematics Department.

Associate Professor Gideon Peyser, Pratt Institute, has been promoted to Professor.

Dr. A. S. Rehm, North American Aviation, Downey, California, has been appointed Assistant Professor at Clarkson College of Technology.

Assistant Professor Wilbur Richardson, Southern Colorado State College, has been promoted to Associate Professor.

Dr. J. A. Roberts, North Carolina State University at Raleigh, has been appointed Assistant Professor at Davidson College.

Professor W. A. Rutledge, University of Tulsa, has been appointed Head of the Department of Mathematics.

Miss Joan Schumaker, Cornell University, has been appointed Assistant Professor at State University College at Geneseo.

Mr. J. A. Shneer, Hughes Aircraft Company, Culver City, California, has been appointed Member of the Technical Staff at the Computation and Data Processing Center of the Aerospace Corporation, El Segundo, California.

Sister Mary Petronia, Mount Mary College, has been appointed Chairman of the Department of Mathematics.

Mr. J. E. Strout, University of Illinois, has been appointed Assistant Professor at Southeast Missouri State College.

Mr. J. C. Warren, Jr., University of Texas, has been appointed Assistant Professor and Chairman of the Mathematics Department of the College of Notre Dame, California.

Professor J. J. Barron, Old Dominion College, died on August 13, 1965. He was a member of the Association for 17 years.

Associate Professor L. L. Garner, University of North Carolina, died on June 12, 1965. He was a member of the Association for 25 years.

Professor Emeritus R. K. Morley, Worcester Polytechnic Institute, died on August 3, 1965. He was a charter member of the Association.

COMMITTEE ON SUPPORT OF RESEARCH IN THE MATHEMATICAL SCIENCES

At its annual meeting in March, 1965 the Division of Mathematics of the National Academy of Sciences—National Research Council voted to establish the Committee on Support of Research in the Mathematical Sciences (COSRMS). The Committee was to be charged with preparing (i) a study of the current state of research in the mathematical sciences and of mathematics education at the undergraduate, graduate, and post-doctoral levels; (ii) a study of the current levels and forms of support of mathematical research by federal and private agencies; (iii) an indication of appropriate support in the immediate future if a healthy state of mathematical activity is to be maintained.

At the same time the Division of Mathematics authorized its Chairman, Mark Kac, to nominate the chairman and members of COSRMS. Following preliminary organizational work during the spring and summer, final appointment of the chairman and members was made by the President of the National Academy of Sciences—National Research Council, Frederick Seitz, in September, 1965. These are as follows.

Lipman Bers, Columbia University (Chairman)
T. W. Anderson, Columbia University
R. H. Bing, University of Wisconsin.
H. W. Bode, Bell Telephone Laboratories, Inc.
R. P. Dilworth, California Institute of Technology
G. E. Forsythe, Stanford University
Mark Kac, Rockefeller University
C. C. Lin, Massachusetts Institute of Technology
J. W. Tukey, Princeton University
F. J. Weyl, Office of Naval Research
Hassler Whitney, Institute for Advanced Study
C. N. Yang, Institute for Advanced Study

The membership of the Committee provides authoritative coverage of various branches of the mathematical community and includes people involved in the applications of mathematics to other sciences.

In addition, COSRMS is served by four panels: a panel on graduate mathematics education under the chairmanship of R. P. Boas of Northwestern University, a panel on undergraduate mathematics education under the chairmanship of John Kemeny of Dartmouth College, a panel on levels and forms of support of mathematical research under the chairmanship of Mina Rees of City University of New York, and a panel on new mathematical centers under the chairmanship of Mark Kac of the Rockefeller University.

COSRMS is closely cooperating with the recently organized Survey Committee of the Conference Board of the Mathematical Sciences under the chairmanship of Gail S. Young of Tulane University.

COSRMS has its headquarters at Columbia University where it held its first meeting on October 8–9, 1965. It is hoped that the major portion of the work of the Committee will be completed within 12 months, though an additional 6 months may be required to prepare the publication of the results. COSRMS will not enter into the important problems of curriculum reforms on the secondary and elementary school level, except insofar as these matters affect the teaching assignments of the colleges or the general research picture.

Clearly the value and success of COSRMS will depend on the full support of the entire mathematical community. Every effort will be made that the final report of the Committee represent, as far as possible, the common opinions of the whole mathematical community and also reflect the opinions of the scientific users of mathematics. The Committee earnestly solicits the ideas, suggestions, and opinions of all of these. Communications to the Committee may be sent to its Executive Director, Truman Botts, at the Department of Mathematics, Columbia University, New York, N. Y. 10027.

MATHEMATICAL ASSOCIATION OF AMERICA

CALENDAR OF FUTURE MEETINGS

Forty-ninth Annual Meeting, Sherman House, Chicago, Illinois, January 26–28, 1966.

Forty-seventh Summer Meeting, Rutgers, The State University, New Brunswick, New Jersey, August 29–31, 1966.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN, Waynesburg College, Waynesburg, Pennsylvania, April 30, 1966.

ILLINOIS, Saint Dominic College, St. Charles, May 13–14, 1966.

INDIANA, Indiana State University, Terre Haute, May 14, 1966.

IOWA, Central College, Pella, April 15, 1966.

KANSAS, University of Kansas, Lawrence, March 26, 1966.

KENTUCKY, University of Kentucky, Lexington, March 26, 1966.

LOUISIANA-MISSISSIPPI, Louisiana State University, Baton Rouge, February 18–19, 1966.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, College of William and Mary, Williamsburg, Va., April 30, 1966.

METROPOLITAN NEW YORK, St. John's University, Jamaica, N. Y., May 7, 1966.

MICHIGAN, Wayne State University, Detroit, April 2, 1966.

MINNESOTA, Macalester College, St. Paul, May, 1966.

MISSOURI, University of Missouri at Rolla, Rolla, April 30, 1966.

NEBRASKA, Nebraska Center for Continuing Education, Lincoln, April 29–30, 1966.

NEW JERSEY, Rutgers, The State University, New Brunswick, November 12, 1966.

NORTHEASTERN

NORTHERN CALIFORNIA, University of California, Berkeley, February 5, 1966.

OHIO, Ohio Wesleyan University, Delaware, April 23, 1966.

OKLAHOMA-ARKANSAS, Oklahoma Baptist University, Shawnee, April 1–2, 1966.

PACIFIC NORTHWEST, University of Victoria, Victoria, British Columbia, June 17, 1966.

PHILADELPHIA

ROCKY MOUNTAIN, Colorado State University, Fort Collins, May 13–14, 1966.

SOUTHEASTERN, Emory University, Atlanta, Georgia, March 25–26, 1966.

SOUTHERN CALIFORNIA, Occidental College, Los Angeles, March 12, 1966.

SOUTHWESTERN, University of New Mexico, Albuquerque, April 1–2, 1966.

TEXAS, Southern Methodist University, Dallas, April 15–16, 1966.

UPPER NEW YORK STATE, St. Bonaventure University, Olean, May 14, 1966.

WISCONSIN, Wisconsin State University, Eau Claire, May 7, 1966.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, Chicago, Illinois, January 24–28, 1966.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Washington State University, Pullman, June 20–24, 1966.

ASSOCIATION FOR COMPUTING MACHINERY, Ambassador Hotel, Los Angeles, August 30–September 1, 1966.

ASSOCIATION FOR SYMBOLIC LOGIC

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

INSTITUTE OF MATHEMATICAL STATISTICS, Rutgers, The State University, New Brunswick, August 30–September 2, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Americana Hotel, New York City, April 13–16, 1966.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Miramar Hotel, Santa Monica, California, May 18–20, 1966.

PI MU EPSILON

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Iowa, Iowa City, May 11–14, 1966.
(Symposium on Numerical Analysis.)

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 2

CONTENTS

FEBRUARY 1966

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE TO AUTHORS

More than enough Main Articles, Mathematical Notes, and Classroom Notes have been accepted (or are under consideration) to fill these Sections of the MONTHLY through the December, 1966 issue. Please do not submit any papers for these Sections until after July 1. The June-July issue will contain addresses to which papers may then be submitted.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

FLOW NETWORKS AND COMBINATORIAL OPERATIONS RESEARCH

D. R. FULKERSON, RAND, Santa Monica, California

PART I: FLOWS IN NETWORKS

1. Maximal flow. A *directed network* (graph) $G=[N;\ \mathcal{A}]$ consists of a finite collection N of elements $1, 2, \cdots, n$ together with a subset $\mathcal{A}$ of the ordered pairs (i, j) of distinct elements of N. The elements of N will be called *nodes*; members of $\mathcal{A}$ are *arcs*. Figure 1.1 shows a directed network having four nodes and six arcs $(1, 2)$, $(1, 3)$, $(2, 3)$, $(2, 4)$, $(3, 2)$, and $(3, 4)$.

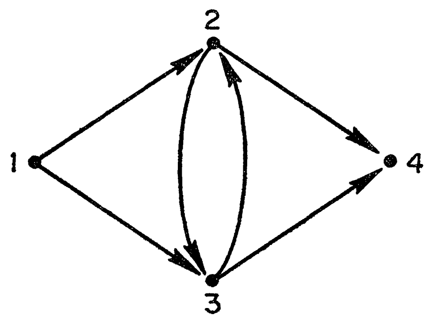


FIG. 1.1

Sometimes we shall also consider *undirected networks*, for which the set $\mathcal{A}$ consists of unordered pairs of nodes. For emphasis, these will then be termed *links*.

Suppose that each arc (i, j) of a directed network has associated with it a nonnegative number c_{ij}, the *capacity* of (i, j), to be thought of as representing the maximal amount of some commodity that can arrive at j from i along (i, j) per unit time in a steady-state situation. Then a natural question is: What is the maximal amount of commodity-flow from some node to another via the entire network? (For example, one might think of a network of city streets, the commodity being cars, and ask for a maximal traffic flow from some point to another.) We may formulate the question mathematically as follows. Let 1 and n be the two nodes in question. A *flow, of amount* v, from 1 *to* n in $G=[N;\ \mathcal{A}]$ is a function x from $\mathcal{A}$ to real numbers (a vector x having components x_{ij} for(i,j) in $\mathcal{A}$) that satisfies the linear equations and inequalities

$$\sum_j x_{ij} - \sum_j x_{ji} = \begin{cases} v, & i = 1, \\ -v, & i = n, \\ 0, & \text{otherwise}, \end{cases} \tag{1.1}$$

$$0 \leqq x_{ij} \leqq c_{ij}, \qquad (i, j) \text{ in } \mathcal{A}. \tag{1.2}$$

In (1.1) the sums are of course over those nodes for which x is defined. We call 1 the *source*, n the *sink*. A *maximal flow* from source to sink is one that maximizes the variable v subject to (1.1), (1.2).

Figure 1.2 shows a flow from source node 1 to sink node 6 of amount 7. In Figure 1.2, the first number of each pair beside an arc is the arc capacity, the second number the arc flow.

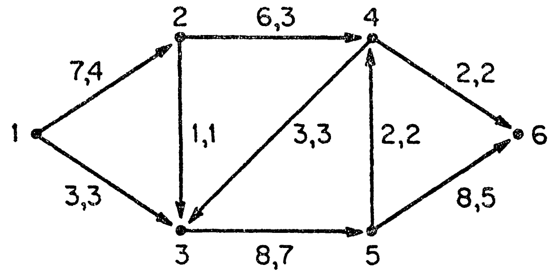


FIG. 1.2

To state the fundamental theorem about maximal flow, we need one other notion, that of a cut. A *cut separating* 1 and n is a partition of the nodes into two complementary sets, I and J, with 1 in I, say, and n in J. The *capacity* of the cut is then

$$\sum_{\substack{i \text{ in } I\\ j \text{ in } J}} c_{ij}. \tag{1.3}$$

(For instance, if $I=\{1, 3, 4\}$ in Fig. 1.2, the cut has capacity $c_{12}+c_{35}+c_{46}=17$.) A cut separating source and sink of minimum capacity is a *minimal* cut, relative to the given source and sink.

Summing the equations (1.1) over i in the source-set I of a cut and using (1.2) shows that

$$v = \sum_{\substack{i \text{ in } I\\ j \text{ in } J}} (x_{ij} - x_{ji}) \leqq \sum_{\substack{i \text{ in } I\\ j \text{ in } J}} c_{ij}. \tag{1.4}$$

In words, for an arbitrary flow and arbitrary cut, the net flow across the cut is the flow amount v, which is consequently bounded above by the cut capacity. Theorem 1.1 below asserts that equality holds in (1.4) for some flow and some cut, and hence the flow is maximal, the cut minimal [11].

THEOREM 1.1. *For any network the maximal flow amount from source to sink is equal to the minimal cut capacity relative to the source and sink.*

Theorem 1.1 is a kind of combinatorial counterpart, for the special case of the maximal flow problem, of the duality theorem for linear programs, and can be deduced from it [5]. But the most revealing proof of Theorem 1.1 uses a simple "marking" or "labeling" process [12] for constructing a maximal flow, which also yields the following theorem.

THEOREM 1.2. *A flow x from source to sink is maximal if and only if there is no flow-augmenting path with respect to x.*

Here we need to say what an x-augmenting path is. First of all, a path from one node to another is a sequence of distinct end-to-end arcs that starts at the first node and terminates at the second; arcs traversed with their direction in going along the path are *forward* arcs of the path, while arcs traversed against their direction are *reverse* arcs of the path. A path from source to sink is x-augmenting provided that $x<c$ on forward arcs and $x>0$ on reverse arcs. For example, the path (1, 2), (2, 4), (5, 4), (5, 6) in Figure 1.2 is an augmenting path for the flow shown there. Figure 1.3 below indicates how such a path can be used to increase the amount of flow from source to sink.

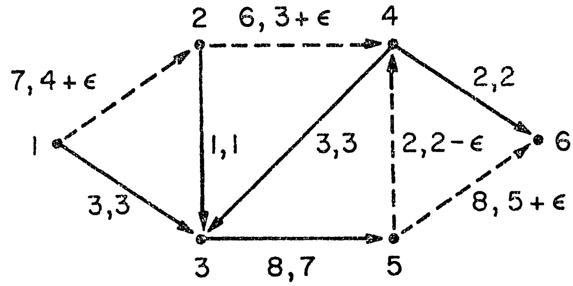


FIG. 1.3

Taking the flow change ϵ along the path as large as possible in Figure 1.3, namely $\epsilon=2$, produces a maximal flow, since the cut $I=\{1, 2, 4\}$, $J=\{3, 5, 6\}$ is then "saturated."

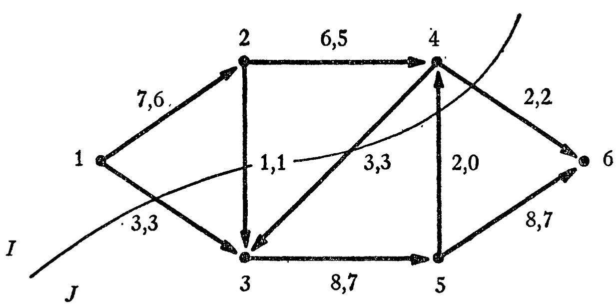


FIG. 1.4

The labeling process of [12] is a systematic and efficient search, fanning out from the source, for a flow augmenting path. If none such exists, the process ends by locating a minimal cut.

The following theorem, of special significance for combinatorial applications, is also a consequence of the procedure sketched above for constructing maximal flow.

THEOREM 1.3. *If all arc capacities are integers, there is an integral maximal flow.*

It is sometimes convenient to alter the constraints (1.2) of the maximal flow problem to

$$l_{ij} \leqq x_{ij} \leqq c_{ij}. \tag{1.5}$$

Here l is a given lower bound function satisfying $l \leqq c$. The analogue of Theorem 1.1 is then

THEOREM 1.4. *If there is a function x satisfying* (1.1) *and* (1.5) *for some number v, then the maximum v subject to these constraints is equal to the minimum of*

$$\sum_{\substack{i \text{ in } I \\ j \text{ in } J}} (c_{ij} - l_{ji}) \tag{1.6}$$

taken over all cuts I, J separating source and sink. On the other hand, the minimum v is equal to the maximum of

$$\sum_{\substack{i \text{ in } I \\ j \text{ in } J}} (l_{ij} - c_{ji}) \tag{1.7}$$

taken over all cuts I, J separating source and sink.

The question of the existence of such a flow x, together with another flow feasibility question, will be discussed in the next section.

2. Feasibility theorems. The constraints of the maximal flow problem are of course always feasible, since $x=0$ satisfies (1.1), (1.2) with $v=0$. By changing the constraints in various ways, interesting feasibility questions arise. Here we shall consider two such, one involving supplies and demands at nodes, the other lower bounds on arc flows, as in (1.5).

Let $G=[N;\,\mathcal{A}]$ have capacity function c, and let S and T be disjoint subsets of N. With each i in S associate a *supply* $a_i \geqq 0$, with each i in T a *demand* $b_i \geqq 0$, and impose the constraints

$$\begin{aligned} \sum_j (x_{ij} - x_{ji}) &\leqq a_i, \quad && i \text{ in } S, \\ &\leqq -b_i, && i \text{ in } T, \\ &= 0, && \text{otherwise}, \end{aligned} \tag{2.1}$$

$$0 \leqq x_{ij} \leqq c_{ij}, \qquad (i, j) \text{ in } \mathcal{A}. \tag{2.2}$$

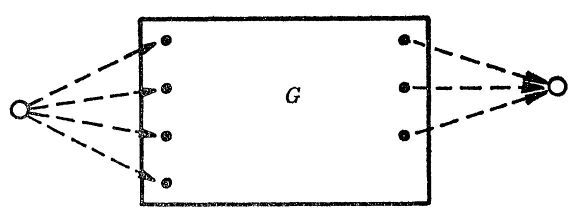


FIG. 2.1

In words, the net flow out of i in S is bounded above by the supply a_i, and the net flow into i in T is bounded below by the demand b_i. When are the supply-demand constraints (2.1), (2.2) feasible?

This question is easily answered by applying Theorem 1.1 to an enlarged network. Extend $G=[N;\ \mathcal{A}]$ to $G^*=[N^*;\ \mathcal{A}^*]$ by adjoining a source 0 and sink $n+1$, together with source arcs $(0, j)$ for j in S, and sink arcs $(i, n+1)$ for i in T. (See Figure 2.1.) The capacity function c^* on $\mathcal{A}^*$ is defined by $c^*_{0,j}=a_j$ for j in S, $c^*_{i,n+1}=b_i$ for i in T, $c^*_{ij}=c_{ij}$ for (i, j) in $\mathcal{A}$. The constraints (2.1) and (2.2) are feasible if and only if the maximal flow amount from source to sink in the enlarged network is at least $\sum_{i \text{ in } T} b_i$, that is, if and only if a maximal flow saturates all sink arcs. Hence we need only construct a maximal flow in order to check the feasibility of (2.1), (2.2). By pushing the analysis a little further, using Theorem 1.1, the following theorem emerges [**21**].

THEOREM 2.1. *The supply-demand constraints* (2.1), (2.2) *are feasible if and only if, for each subset T' of T, there is a flow $x(T')$ that satisfies the aggregate demand $\sum_{i \text{ in } T'} b_i$ without violating the supply limitations at nodes of S.*

Here satisfying the aggregate demand over T' means that the net flow into the set T' must be at least $\sum_{i \text{ in } T'} b_i$, without regard for the individual demands in T'. The necessity of the condition is of course clear; sufficiency asserts the existence of a single flow x meeting all individual demands provided the flows $x(T')$ exist for all subsets T' of T.

It should be noted that if the functions a, b, c of (2.1), (2.2) are integral valued, and if feasible flows exist, then there is an integral feasible flow. This follows from Theorem 1.3 and the conversion of (2.1), (2.2) to a maximal flow problem. A similar integrity statement holds for the situation of Theorem 1.4, and indeed, for all the flow problems to be discussed in any detail in this survey.

We turn now to a consideration of lower bounds on arc flows, as in (1.5), and pose the resulting feasibility question in terms of circulations, i.e., flows that are source-sink free, instead of flows from source to sink. (One can always add a "return-flow" arc from sink to source to convert to circulations.) Thus we are questioning the feasibility of the constraints

$$\sum_j (x_{ij} - x_{ji}) = 0, \qquad i \text{ in } N, \tag{2.3}$$

$$l_{ij} \leqq x_{ij} \leqq c_{ij}, \qquad (i, j) \text{ in } \mathcal{A}. \tag{2.4}$$

The following theorem answers the question [**26**]. Its proof can be made to rely on Theorem 1.1 [**15**].

THEOREM 2.2. *The constraints* (2.3), (2.4) *are feasible if and only if*

$$\sum_{\substack{i \text{ in } I \\ j \text{ in } J}} c_{ij} \geqq \sum_{\substack{i \text{ in } I \\ j \text{ in } J}} l_{ji} \tag{2.5}$$

holds for all partitions I, J of N.

Again the necessity is clear, since (2.5) simply says there must be sufficient escape capacity from the set I to take care of the flow forced into I by the function l. But sufficiency is not obvious.

Other useful flow feasibility theorems have been deduced [**19**, **26**]. In each case Theorem 1.1 can be used as the main tool in a proof.

3. Minimal cost flows. One of the most practical problem areas involving network flows is that of constructing flows satisfying constraints of various kinds and minimizing cost. The standard linear programming transportation problem, which has an extensive literature, is in this category.

We put the problem as follows. Each arc (i, j) of a network $G=[N; \mathcal{A}]$ has a capacity c_{ij} and a cost a_{ij}. It is desired to construct a flow x from source to sink of specified amount v that minimizes the total flow cost

$$\sum_{(i,j) \text{ in } \mathcal{A}} a_{ij}x_{ij} \tag{3.1}$$

over all flows that send v units from source to sink. In many applications one has supplies of a commodity at certain points in a transportation network, demands at others, and the objective is to satisfy the demands from the supplies at minimum cost.

By treating v as a parameter, the method for constructing maximal flows can be used to construct minimal cost flows throughout the feasible range of v. Indeed, the solution procedure can be viewed as one of solving a sequence of maximal flow problems, each on a subnetwork of the original one [**14**]. Another, not essentially different, viewpoint is provided by the following theorem [**1**, **29**].

THEOREM 3.1. *Let x be a minimal cost flow from source to sink of amount v. Then the flow obtained from x by adding $\epsilon>0$ to the flow in forward arcs of a minimal cost x-augmenting path, and subtracting ϵ from the flow in reverse arcs of this path, is a minimal cost flow of amount $v+\epsilon$.*

Here the cost of a path is the sum of arc costs over forward arcs minus the corresponding sum over reverse arcs, i.e., the cost of "sending an additional unit" via the path.

Thus, if all arc costs a_{ij} are nonnegative, for example, one can start with the zero flow and apply Theorem 3.1 to obtain minimal cost flows for increasing v. (The cost profile thereby generated is piecewise linear and convex.) All that is needed to make this an explicit algorithm is a method of searching for a minimal cost flow augmenting path. Various ways of doing this can be described. One such will be given in Part II, section 1.

Another method [**17**] for constructing minimal cost flows poses the problem in circulation format, that is, (3.1) is to be minimized subject to (2.3), (2.4). This construction has a number of advantages, principally in terms of generality and flexibility. For instance, it may be started with any circulation; even (2.4) need not be satisfied initially. Also, no assumption about the cost function is required.

These methods produce integral flows in case the arc capacities (and lower bounds) are integers. Theoretical upper bounds on the computing task, ones that are quite good, are easily obtained in each case. This may be contrasted with the situation for general linear programs, where decent upper bounds on solution methods are unknown.

4. Maximal dynamic flow. Suppose that each arc (i, j) of a network G has not only a capacity, but a transit time t_{ij} as well, and that we are interested in determining the maximum amount of flow that can reach sink n from source 1 in a specified number t of time periods. This dynamic flow problem can always be treated as a static flow problem in a time-expanded version G_t of G. For example, if the given network G is that of Figure 4.1 and if each arc of G has unit transit time, then G_3 is shown in Figure 4.2. (We have included "storage arcs" leading from a location to itself one unit of time later.)

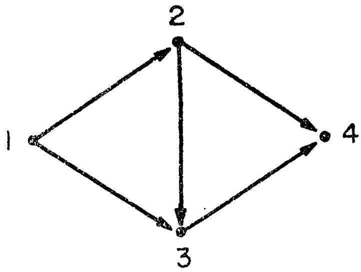


FIG. 4.1

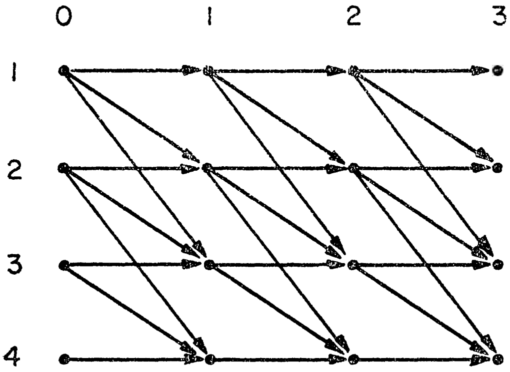


FIG. 4.2

Expanding the network in this way puts one back in the static case. Moreover, arc capacities and transit times can vary with time and this is still so. However, if each capacity and transit time is fixed over time, the problem can be solved in the smaller network G. Specifically, a maximal dynamic flow for t periods can be generated from a static flow x in G of amount v that minimizes the linear form

$$\sum_{(i,j) \text{ in } \mathfrak{A}} t_{ij}x_{ij} - (t+1)v \tag{4.1}$$

over all flows in G from source 1 to sink n [14]. By adding the return-flow arc $(n, 1)$ to G with $c_{n1} = \infty$, $t_{n1} = -(t+1)$, the problem may be viewed as one of constructing a circulation that minimizes the "cost" form (4.1).

5. Multi-terminal maximal flow. Heretofore we have phrased statements in terms of directed networks. In this section we confine the discussion to undirected flow networks, by which we merely mean the following. A link (i, j) can carry flow in either direction and has the same flow capacity each way. Thus

one can think of an arc (i, j) with capacity c_{ij} and an arc (j, i) with capacity $c_{ji}=c_{ij}$. The assumption of a symmetric capacity function c makes the results described in this section considerably simpler and more appealing than they would otherwise be.

Instead of dealing with a single source and sink, we shift attention to all pairs of distinct nodes taken as terminals for flows. These flows are not to be thought of as occurring simultaneously.

Let v_{ij} denote the maximal flow amount from i to j. Thus the function v is symmetric, $v_{ij}=v_{ji}$, and may be determined explicitly for an n-node network by solving $n(n-1)/2$ maximal flow problems. There is, however, a much simpler way of determining the function v, one that involves the solution of only $n-1$ maximal flow problems; in addition, there is a simple condition in order that a symmetric function v be realizable as the maximal multiterminal flow function of some undirected network [22].

THEOREM 5.1. *A symmetric, nonnegative function v is realizable by an undirected network if and only if v satisfies*

$$v_{ij} \geqq \min (v_{ik}, v_{kj}) \tag{5.1}$$

for all triples i, j, k.

The necessity of the "triangle inequality" (5.1) follows easily from Theorem 1.1.

The condition (5.1) imposes severe limitations on the function v. For instance, among the three functional values appearing in (5.1), two must be equal and the third no smaller than their common value. It also follows that if the network has n nodes, v can take on at most $n-1$ numerically different functional values. It is not altogether surprising, therefore, that v can be determined by a

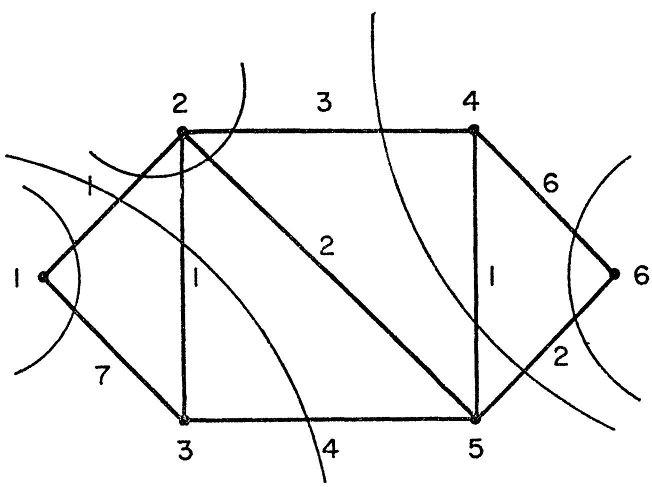


FIG. 5.1

simpler process than solving all single-terminal maximal flow problems. This process systematically picks out precisely $n-1$ cuts in the network having the property that v_{ij} is determined by the minimum one of these cuts separating i and j [22]. For example, in the network of Figure 5.1 the relevant cuts are those shown. Thus, for instance, since nodes 1 and 4 are separated by the three cuts (1/2, 3, 4, 5, 6), (1, 3/2, 4, 5, 6), (1, 2, 3, 5/4, 6) having capacities 8, 6, 6 respectively, then $v_{14}=\min(8, 6, 6)=6$.

6. Other flow problems. The flow problems that have been discussed thus far all have the useful and pleasant feature that the assumption of integral data implies the existence of an integral solution. A number of flow problems that do not share this property have also been studied. Among these we mention flows in networks with gains [29], simultaneous multi-terminal flows [13], and problems involving optimal synthesis of flow networks that meet specified requirements [16, 22]. Methods of solution for such problems are, in general, more complicated than methods for those we have discussed. One rather surprising exception to this statement is the following synthesis problem. Suppose it is desired to construct an undirected network on a specified number of nodes so that $v_{ij} \geqq r_{ij}$ for stipulated requirements r_{ij}, with the total sum of link capacities of the network minimal. A very simple combinatorial method of solution for this synthesis problem is given in [22].

PART II: COMBINATORIAL PROBLEMS

1. Network potentials and shortest chains. Consider a directed network in which each arc (i, j) has associated with it a positive number a_{ij}, which may be thought of as the length of the arc, or the cost of traversing the arc. How does one determine a shortest chain from some node to another? Here we have used *chain* to mean a path containing only forward arcs, the length of the chain being obtained by adding its arc lengths.

While this is a purely combinatorial problem, it may also be viewed as a flow problem simply by imposing a cost a_{ij} per unit flow in (i, j), taking all arc capacities infinite, and asking for a minimal cost flow of one unit from the first node to the second. An integral optimal flow corresponding to $v=1$ singles out a shortest chain.

Many ways of locating shortest chains efficiently have been suggested. We describe one [10]. Like others, it simultaneously finds shortest chains from the first node to all others reachable by chains.

In this method each node i will initially be assigned a number π_i. These node numbers, which we shall refer to as *potentials*, will then be revised in an iterative fashion. Let 1 be the first node. To start, take $\pi_1=0$, $\pi_i=\infty$ for $i\neq 1$. Then search the list of arcs for an arc (i, j) whose end potentials satisfy

$$\pi_i + a_{ij} < \pi_j. \tag{1.1}$$

(Here $\infty + a = \infty$.) If such an arc is found, change π_j to $\pi_j' = \pi_i + a_{ij}$, and search

again for an arc satisfying (1.1), using the new node potentials. Stop the process when the node potentials satisfy

$$\pi_i + a_{ij} \geqq \pi_j \tag{1.2}$$

for all arcs.

It is not hard to show that the process terminates, and that when this happens, the potential π_j is the length of a shortest chain from 1 to j. (Here $\pi_j = \infty$ at termination means there is no chain from 1 to j.) A shortest chain from 1 to j can be found by tracing back from j to 1 along arcs satisfying (1.2) with equality (see Figure 1.1).

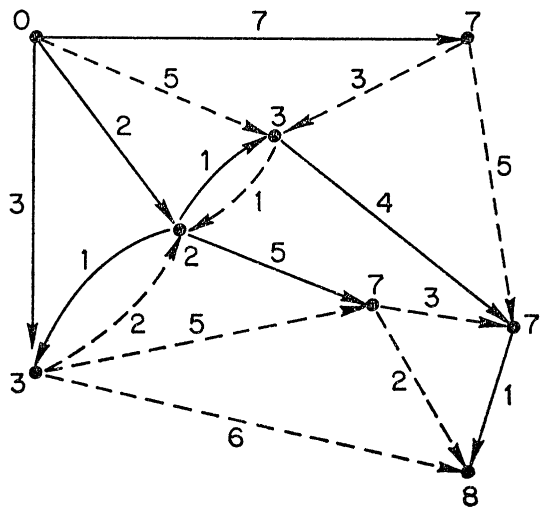


Fig. 1.1

Practical applications that require shortest chains are numerous. For instance, in making up a table of highway distances between cities, a shortest chain between each pair needs to be found. A less obvious application is the discrete version of the problem of determining the least time for an airplane to climb to a given altitude [2]. Some other applications will be discussed in following sections.

While we have assumed positive lengths for the method described above, this assumption can be weakened. Call a chain of arcs leading from a node to itself a *directed cycle*. Then it is enough to suppose that all directed cycle lengths are nonnegative.

If directed cycle costs are nonnegative, the minimum cost flow problem of Part I, section 3, can be solved by repeatedly finding cheapest chains in suitable networks. Because of the assumption on the cost function a, we may start with the zero flow. Thus, using Theorem 3.1, it is enough to reduce the problem of finding a cheapest flow augmenting path with respect to a minimal cost flow x of amount v to that of finding a cheapest chain. Define a new network $G' = [N; \mathcal{A}']$

from the given one $G=[N;\ \mathcal{A}]$ and the flow x as follows. First note that we may assume $x_{ij}\cdot x_{ji}=0$, since $a_{ij}+a_{ji}\geqq 0$. Now put $(i,\ j)$ in $\mathcal{A}'$ if either $x_{ij}<c_{ij}$ or $x_{ji}>0$ and define a' by

$$a'_{ij} = \begin{cases} a_{ij} & \text{if} \quad x_{ij} < c_{ij} \quad \text{and} \quad x_{ji} = 0, \\ -a_{ji} & \text{if} \quad x_{ji} > 0. \end{cases} \tag{1.3}$$

Thus a chain from source to sink in the new network corresponds to an x-augmenting path in the old, and these have the same cost. Moreover, since x is a minimal cost flow, the function a' satisfies the nonnegative directed cycle condition. Hence the method of this section can be used to construct minimal cost flows of successively larger amounts.

2. Optimal chains in acyclic networks. If the network is acyclic (contains no directed cycles), the shortest chain method of the last section can be modified in such a way that, once a potential is assigned a node, it remains unchanged. One can begin by numbering the nodes so that if $(i,\ j)$ is an arc, then $i<j$. Such a numbering can be obtained as follows. Since the network is acyclic, there are nodes having only outward-pointing arcs. Number these nodes $1,\ 2,\ \cdots,\ k$ in any order. Next delete these nodes and all their arcs, search the new network for nodes having only outward-pointing arcs, and number these, starting with $k+1$. Repetition of this process leads to the desired kind of numbering (see Figure 2.1).

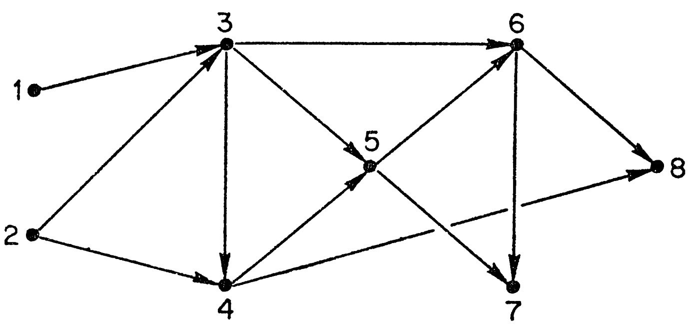


FIG. 2.1

If we wish to find shortest chains from node k to all other nodes reachable from k by chains, the calculation is now trivial. Simply define $\pi_k,\ \pi_{k+1},\ \cdots,\ \pi_n$ recursively by

$$\begin{cases} \pi_k = 0 \\ \cdots \\ \pi_j = \min\limits_{k\leq i<j} (\pi_i + a_{ij}), \quad j = k+1,\ \cdots,\ n. \end{cases} \tag{2.1}$$

Here the minimum is of course taken over i such that $(i,\ j)$ is an arc.

Longest chains in acyclic networks can be computed by replacing "min" by "max" in (2.1).

The recursion (2.1), of dynamic programming type, can be applied in a number of problems. We shall discuss three such applications in the following sections.

3. The knapsack problem. Suppose there are K objects, the i-th object having weight w_i and value v_i, and that it is desired to find the most valuable subset of objects whose total weight does not exceed W. Thus we wish to maximize

$$\sum_{i \text{ in } S} v_i \tag{3.1}$$

over subsets $S \subseteq \{1, 2, \cdots, K\}$ such that

$$\sum_{i \text{ in } S} w_i \leqq W. \tag{3.2}$$

We take $w_1, w_2, \cdots, w_K, W$ to be positive integers.

This combinatorial problem, commonly referred to as the knapsack problem, can be viewed as one of finding a longest chain in a suitable acyclic network. Let the network have nodes denoted by ordered pairs (i, w), $i=0, 1, \cdots, K$, $w=0, 1, \cdots, W$. The node (i, w) has two arcs leading into it, one from $(i-1, w)$, the other from $(i-1, w-w_i)$, provided these exist. (See Figure 3.1.) The length of the first arc is zero; the other has length v_i. In addition we put in a starting node and join it to all of the nodes $(0, w)$ by arcs of length zero. Then chains from the starting node to (i, w) correspond to subsets of the first i objects whose total weight is at most w, the length of the chain being the value of the subset.

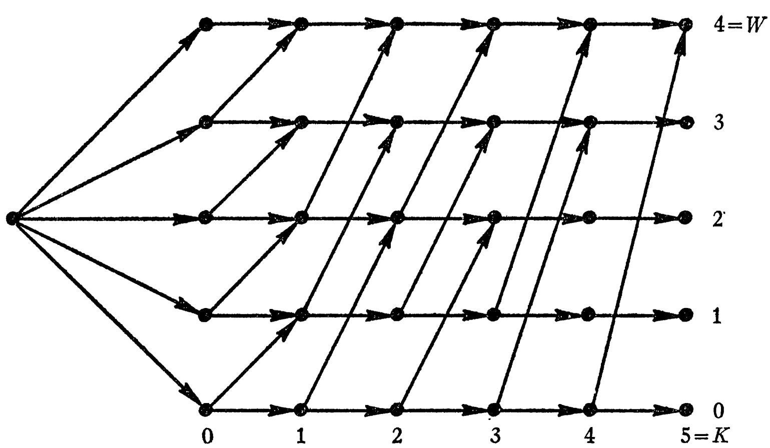


FIG. 3.1

4. Equipment replacement. As equipment deteriorates with age, and improved equipment becomes available on the market, a time may be reached

when the purchase cost of new equipment is repaid by its potential future earnings. One is then faced with the problem of determining an optimal replacement policy [9].

For simplicity, consider a single machine and suppose that at the beginning of each of K periods of time it must be decided whether to keep the machine another period or purchase a new one. Let $r(i, t)$ denote the revenue obtainable during period i from a machine which starts the period at age t (the function r may reflect upkeep costs), and let $c(i, t)$ denote the cost of replacing a machine of age t with a new machine if the replacement occurs at the beginning of period i. Thus replacing a machine of age t at period i gives a net return for the period of $r(i, 0)-c(i, t)$.

The acyclic network shown in Figure 4.1 indicates one formulation in terms of chains. Again nodes are points (i, t), $i=0, 1, \cdots, K$, $t=0, 1, \cdots, T$. (Here T is some sufficiently large integer; if we start with a machine of age t, $T=K+t$ will do.) In general, two arcs, reflecting the possibilities of keeping or replacing, lead from (i, t), the "keep" arc going to $(i+1, t+1)$, the "replace" arc to $(i+1, 1)$. The first of these has length $r(i, t)$, the second has length $r(i, 0)-c(i, t)$. We may also put in a sink node with arcs of length zero leading into it from the nodes (K, t), $t=0, 1, \cdots, T$. Then chains from $(0, t)$ to the sink correspond to the various replacement policies starting with a machine of age t, the length of the chain being the total return from the policy.

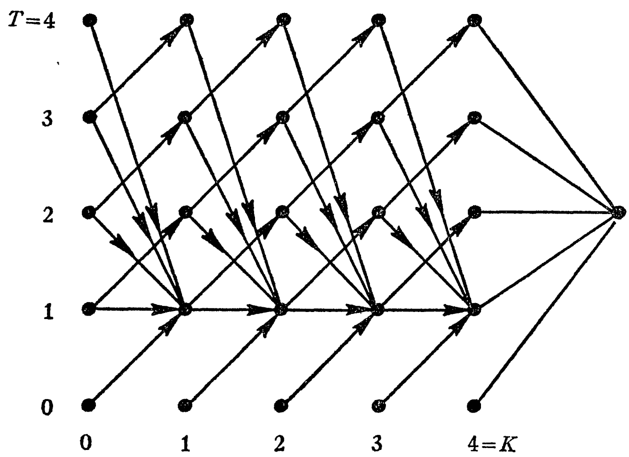


FIG. 4.1

A simpler network for this problem is shown in Figure 4.2. In Figure 4.2 an arc (i, j) corresponds to keeping the machine throughout periods $i, i+1, \cdots, j-1$ and replacing it at the start of period j, the associated length being the return obtained from this action. Thus a longest chain from source 1 to sink K is to be found.

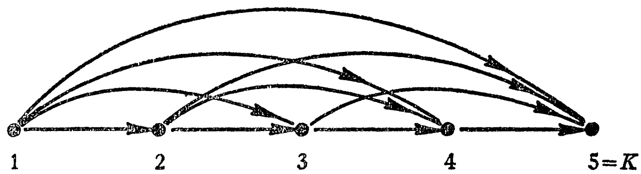


FIG. 4.2

The examples of this section and the preceding section are typical discrete dynamic programming problems. Such problems can always be viewed as seeking optimal chains in appropriate acyclic networks.

5. Project planning. One of the most popular combinatorial applications involving networks deals with the planning and scheduling of large, complicated projects [33]. Suppose that a project of some kind (the construction of a bridge, for example) is broken down into many individual jobs. Certain of these jobs will have to be finished before others can be started. We may depict the order relations among the jobs by means of an acyclic network whose arcs represent jobs. To take a simple case, suppose there are five jobs with the ordering: 1 precedes 3; 1 and 2 precede 4; 1, 2, 3 and 4 precede 5. This may be pictured by the network shown in Figure 5.1.

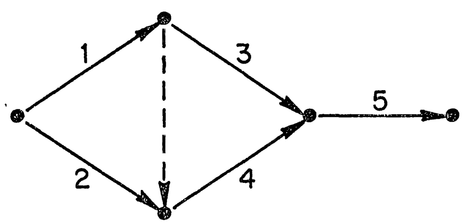


FIG. 5.1

Notice that we have added a "dummy" job, the dotted arc of Figure 5.1, to maintain the proper order relations among the jobs. The use of dummies permits a network representation of this kind for any project (finite partially ordered set).

Assuming that each job has a known duration time (dummies have zero duration time), that the only scheduling restriction is that all inward-pointing jobs at a node must be finished before any outward-pointing job can be started, it follows that the minimum time to complete the entire project is equal to the length of a longest chain of jobs. Hence the minimum project duration time can be calculated easily.

Although a fixed time has been assumed for each job, it may be the case that by spending more money, a job can be expedited. The question then arises: Which jobs should money be spent on and how much, in order that the project be finished by a given date at minimum cost? If the time-cost relation for each

job is linear, this problem can be shown to be a minimal cost flow problem of the kind described in Part I, section 3 [**18, 33**].

6. Minimal chain coverings of acyclic networks. The following question concerning acyclic networks has both theoretical and practical interest: What is the minimum number of chains required to cover a given subset of arcs? We first show how flows may be used to answer this question, and then give a practical interpretation.

Let the subset of arcs be denoted by $\mathcal{A}'$. To rephrase the question, we seek the minimum number of chains in the acyclic network G such that every arc of $\mathcal{A}'$ belongs to at least one of these chains. Theorems 1.3 and 1.4 can be used to provide an answer to the question, as sources in G take all nodes with only outward-pointing arcs; as sinks take all nodes with only inward-pointing arcs. Now place a lower bound of 1 on flow in arcs of $\mathcal{A}'$, 0 on arcs not in $\mathcal{A}'$, and take all arc capacities infinite. Then an integral flow through G of amount v picks out v chains in G that cover all arcs of $\mathcal{A}'$, and the second half of Theorem 1.4 implies

THEOREM 5.1. *The minimum number of chains in an acyclic network needed to cover a subset of arcs is equal to the maximum number of arcs of the subset having the property that no two belong to any chain.*

Theorem 5.1 is a generalization of a known result on chain decompositions of partially ordered sets [**8**].

A practical instance of this situation arises if we think of an airline, say, attempting to meet a fixed flight schedule with the minimum number of planes, all of the same type [**4**]. Let the individual flights be numbered $1, 2, \cdots, n$. Start and finish times $s_i < f_i$ are known for each flight, and the times t_{ij} to return from the destination of the ith flight to the origin of the jth flight are also known. The flights can be partially ordered by saying that i precedes j if $f_i + t_{ij} \leqq s_j$, and the resulting partially ordered set represented by an acyclic network (as in the preceding section). A chain in the network represents a possible assignment of flights to one aircraft. The problem then is to cover the nondummy arcs (those corresponding to actual flights) with the minimum number of chains. Theorem 5.1 asserts that this number is equal to the maximum number of flights, no two of which can be accomplished by a single plane.

Problems of this nature become considerably more complicated if the assumption of a fixed schedule is dropped. For instance, suppose the times s_i, f_i are at our disposal subject to the restriction that $f_i - s_i = t_i$, with the duration times t_i known, as well as the reassignment times t_{ij}. The problem might then be to arrange a schedule completing all flights by a given time and requiring the minimum number of planes, or to finish all flights at the earliest possible time with a fixed number of planes. For such scheduling problems there is very little known in the way of general theoretical results or good computational procedures. However, some special results have been deduced [**27, 31**].

7. Assignment problems. The following is typical of an important and well-known class of combinatorial problems having network flow formulations. Suppose there are m men and n jobs, and that it is known whether or not man i is qualified to fill job j, $i=1, 2, \cdots, m$, $j=1, 2, \cdots, n$. When is it possible to fill all jobs with qualified men and how does one determine such an assignment?

Using Theorem 1.3, the problem may be phrased in terms of flows. Corresponding to man i take a source node i, to job j a sink node j, and direct an arc from i to j if man i is qualified for job j. (See Figure 7.1.) Impose a demand of 1 unit at each sink and let each source have a supply of 1 unit. All arc capacities may be taken infinite. The problem of assigning men to jobs thus becomes that of constructing a flow (integral, of course) meeting the demands from the supplies.

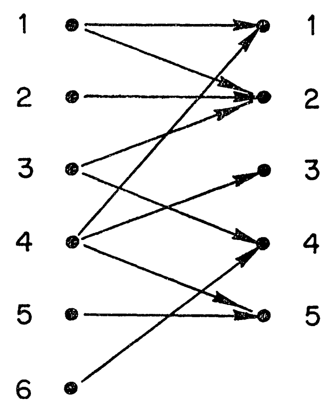


FIG. 7.1

Combinatorial interpretations of Theorems 1.1 and 2.1 for this situation lead in the first instance to a well-known theorem about maximum matchings and minimum covers in bipartite networks [35], and in the second instance to an equally well-known, and equivalent, theorem concerning systems of distinct representatives for subsets of a given set [24].

A more general assignment problem, usually referred to as that of optimal assignment [36, 42], assumes that man i in job j is worth a_{ij} units, and the total worth of an assignment is given by the sum of the numbers a_{ij} taken over the individual man-job matchings in the assignment. The problem then is to construct an assignment of maximal worth. By taking the cost per unit of flow in arc (i, j) to be $-a_{ij}$, the optimal assignment problem is seen to be a special case of the minimal cost flow problem.

More complicated personnel assignment models have been formulated in terms of flow networks. For instance, one which involves the recruiting, training, and retraining of military personnel to meet stipulated requirements in various job specialties over time has been treated in this way [23].

Applications of the optimal assignment model to other kinds of problems

are very numerous. We mention one which involves the optimal depletion of inventory [7]. Suppose a stockpile consists of m items of the same kind, and that the age t_i of item i is known. Also known is a function $u(t)$ giving the utility for an item of age t when withdrawn from the stockpile, together with a schedule of demands specifying the times at which items will be required. The problem is to determine that order of item issue which maximizes the total utility while meeting the demand schedule. (For a concrete example, suppose one has m bottles of wine in his cellar, the ages of each being known, and consumes one bottle of wine weekly. The utility function for wine might appear as in Figure 7.2.) The utility of item i issued at time j is given by $u_{ij}=u(t_i+j)$, and hence the problem is to find an assignment of items to times which is optimal in terms of the u_{ij}.

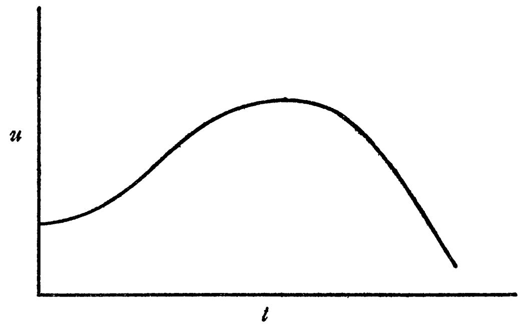


FIG. 7.2

If the utility function is convex, there is a simple rule for solving the problem: Issue the youngest item first, then the next youngest, and so on. This policy, sometimes called LIFO (last in first out), may be shown optimal here by a simple interchange argument. Similarly, if the utility function is concave, the reverse rule FIFO (first in first out) solves the problem. In general, however, no such simple rule works and an optimal assignment needs to be computed.

8. Production and inventory planning. Problems involving dynamic production and inventory programs for a single type item have received considerable study. A very simple deterministic problem in this category is the following. Suppose there are n periods of time with known period demands $b_1, b_2, \cdots, b_n$ for the item, that the unit cost of production in period i is p_i, and the unit cost of storage from period i to $i+1$ is s_i. What pattern of production and storage meets the demands at minimum cost?

The network shown in Figure 8.1 assumes that production in period i can be used to satisfy demand in period i. The ith "production" arc (source arc) has infinite capacity and cost p_i; the ith "storage" arc has infinite capacity and cost s_i; the ith "demand" arc (sink arc) has capacity b_i and zero cost. The problem then is to determine a flow of amount $v=\sum_i b_i$ from source to sink that min-

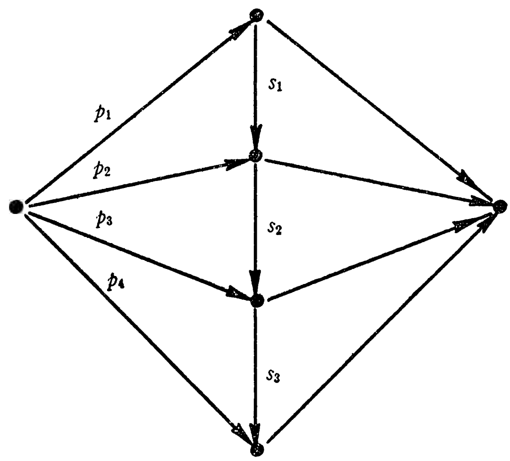


FIG. 8.1

imizes cost. Clearly production and storage capacities may be introduced if desired. But if these are left infinite, there is a very simple rule for solving the problem. For the ith demand, compare the chain costs

$$\begin{cases} p_1 + s_1 + \cdots + s_{i-1} \\ p_2 + s_2 + \cdots + s_{i-1} \\ \cdots \\ p_i \end{cases}$$

and take the smallest of these. Then send b_i units of flow along the corresponding chain. An almost equally simple rule works in case each period's production cost is convex in the number of items produced [30].

If it is assumed that demands do not have to be satisfied, but that unfulfilled demand in period i results in a penalty cost c_i per unit, we may place a flow cost $-c_i$ on the ith sink arc and solve the minimal cost flow problem parametrically in the flow amount v, selecting that v which gives the least cost.

9. Optimal capacity scheduling. The model of this section, proposed and studied in [41], is a rather general one which can be shown to include several of those previously discussed here. One version of the model is described in [41] as follows: "A decision maker must contract for warehousing capacity over n time periods, the minimal capacity requirement for each period being deterministically specified. His economic problem arises because savings may possibly accrue by his undertaking long-term leasing or contracting at favorable periods of time, even though such commitments may necessitate leaving some of the capacity idle during several periods."

To put the problem mathematically, let d_i be the minimal capacity demand in period i. Let x_{ij}, $i<j$, be the number of units of capacity acquired at the begin-

ning of period i, available for possible use during periods $i, i+1, \cdots, j-1$, and relinquished at the beginning of period j, and let a_{ij} be the associated unit cost. Then the problem is to find $x_{ij} \geqq 0$ that minimize

$$\sum_{i=1}^{n} \sum_{j=i+1}^{n+1} a_{ij} x_{ij} \tag{9.1}$$

subject to the constraints

$$\sum_{i=1}^{k} \sum_{j=k+1}^{n+1} x_{ij} \geqq d_k, \qquad k = 1, 2, \cdots, n. \tag{9.2}$$

To see that the constraints (9.2) describe flows, first rewrite (9.2) as

$$\sum_{i=1}^{k} \sum_{j=k+1}^{n+1} x_{ij} - y_k = d_k, \qquad y_k \geqq 0. \tag{9.3}$$

Next successively subtract the $(k-1)$-st equation from the kth, $k=n, n-1, \cdots, 2$, to obtain an equivalent system of constraints. The result is

$$\begin{aligned}
&\sum_{j=2}^{n} x_{ij} - y_1 = d_1, \\
&\cdots \\
&\sum_{j=k+1}^{n} x_{kj} - \sum_{i=1}^{k-1} x_{ik} + y_{k-1} - y_k = d_k - d_{k-1}, \qquad k = 2, \cdots, n, \\
&\cdots \\
&-\sum_{i=1}^{n-1} x_{in} + y_n = - d_n,
\end{aligned} \tag{9.4}$$

subject to which the linear form (9.1) is to be minimized.

The corresponding network is shown in Figure 9.1 below.

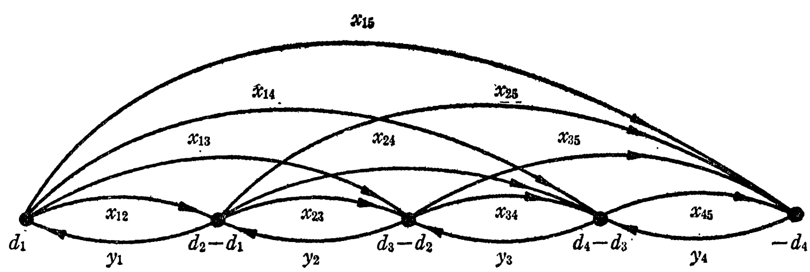


FIG. 9.1

Here x_{ij} is the flow in (i, j) and a_{ij} is the cost per unit of flow; y_i is the flow in $(i+1, i)$, with $a_{i+1,i}=0$. Nodes i for which $d_i-d_{i-1}>0$ are sources with supplies d_i-d_{i-1}; nodes i for which $d_i-d_{i-1}<0$ are sinks with demands $-(d_i-d_{i-1})$.

Referring to Figure 4.2, Part II, and Figure 9.1 above, it is apparent that

the equipment replacement problem can be viewed as a special case of capacity scheduling by taking $d_i=1$, all i.

A number of other situations that can be interpreted in terms of capacity scheduling are described in [41]. Mentioned are models involving checkout and replacement of stochastically failing mechanisms; determination of economic lot sizes, product assortment, and batch queuing policies; labor-force planning; and multi-commodity warehousing decisions.

One application (the dynamic economic lot size model [43]) deals with the problem described in section 8, where production costs now are concave functions of the number of items produced, and demands must be satisfied. Generally speaking, concavity makes minimization problems difficult, but here it can be seen that it is uneconomical to both produce in a period and carry inventory into the period, and hence there is an optimal policy of the following kind. The total time interval is broken into subintervals, with enough production at the beginning of each of these to satisfy its aggregate demand. Thus finding an optimal policy can be formulated in terms of capacity scheduling by letting a_{ij} be the total cost (including storage) associated with producing enough in period i to satisfy the demands for periods $i, i+1, \cdots, j-1$, and by taking all $d_i=1$. In short, the problem has been reduced to one of finding a cheapest chain from source to sink in the network of Figure 9.1.

10. Minimal spanning trees. A network combinatorial problem for which there is a particularly simple solution method is that of selecting a minimum spanning subtree from an undirected network each of whose links has a length or cost. We may illustrate this problem with the following example. Imagine a number n of cities on a map and suppose that the cost of installing a communication link between cities i and j is $a_{ij}=a_{ji}\geqq 0$. Each city must be connected, directly or indirectly, to all others, and this is to be done at minimum total cost. Clearly attention can be confined to trees (acyclic and connected networks of links), for if a connected network contains a cycle, removing one link of the cycle leaves the network connected and reduces cost. A minimal cost tree can be found easily as follows [37]. Select the cheapest link, then the next cheapest, and so on, being sure at each stage that no subset of the selected links forms a cycle. After $n-1$ selections, a cheapest tree has been constructed.

For example, in the network of Figure 10.1, this procedure might lead to the minimal cost tree shown in heavy links.

While it is not difficult to prove that this method solves the problem, it is nonetheless remarkable that being greedy at each stage works. There are few extremal combinatorial problems for which it does.

There is an interesting relation between the minimal spanning tree problem and another, which sounds on the surface to be very different. Think of the network in Figure 10.1 as being a highway map, where the number recorded beside each link is the maximum elevation encountered in traversing the link. Suppose someone who plans to drive from i to j dislikes high altitudes and hence

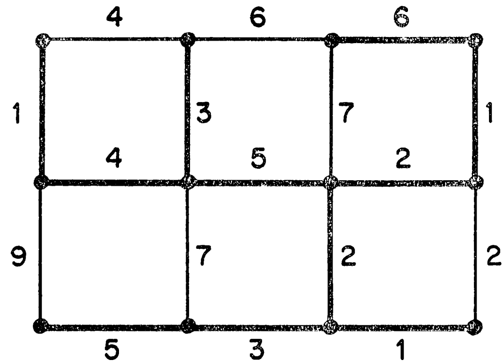


FIG. 10.1

wants to find a path connecting i and j that minimizes the maximum altitude. This problem is related to the shortest chain problem in the sense that methods for solving the latter are easily modified to solve the former, and this is so in either the directed or undirected case [40]. But it is also true in the undirected case that the minimal spanning tree solves the problem, and for all pairs of cities. That is, the unique path in the minimal spanning tree joining a pair of cities minimizes the path height [28]. Here we have used "path height" to mean the maximum number on the path.

There is also a min-max theorem concerning paths and cuts for this problem. Call the minimum link number in a cut the "cut height." Then it may be verified that the minimum height of paths joining two nodes is equal to the maximum height of cuts separating the two.

11. The traveling-salesman problem. Many problems that involve minimal connecting networks, and hence superficially resemble that of the last section, have no known simple solution procedures. For example, consider a network connecting a number of cities, with the length of each link being known, and imagine a traveling salesman who must start at some city, visit each of the others just once, and then return to the starting city. How does the salesman determine an itinerary that minimizes the total distance traveled?

A cycle that passes through every node of a network just once is usually called a Hamiltonian cycle. For brevity, we refer to it as a tour. Thus the problem asks for a shortest tour. (Of course the given network may contain no tour, but the existence question can be subsumed by considering all possible links present, those not corresponding to original ones having very large lengths.)

Not a great deal is known, either theoretically or computationally, about this problem, except that it is hard. On the theoretical side, for instance, there seem to be no simple conditions that are necessary and sufficient for a given network to contain a tour. On the computational side, while many methods for determining a shortest tour have been proposed, it is safe to assert that no one of

these would guarantee that a problem involving 100 cities, say, could be solved in a reasonable length of time.

12. Minimal k-connected networks. In considering the synthesis of reliable communication networks with respect to link failure, the following question may be raised. Suppose given the complete, undirected network G on n nodes, where again each link of G has an associated number, the cost of installing a communication link between its end-nodes. For each $k=1, 2, \cdots, n-1$, find a minimal cost k-link-connected spanning subnetwork of G [20]. Here a k-link-connected network is one in which at least k links must be suppressed in order to disconnect the network. Another way of characterizing this property is to say that every pair of nodes is joined by at least k link-disjoint paths. Thus k might be thought of as the "reliability level" of the communication network, and the practical problem is to minimize cost while achieving a stipulated reliability level.

For $k=1$, the problem is that of section 10, and hence is readily solved. For $k=2$ and all link costs 1 or ∞, the problem becomes that of determining whether a given network (the subnetwork of unit cost links) contains a tour, and is thus already difficult. But if all link costs are equal, the answer is known. Here the problem is to determine the minimum number of links required in a k-link-connected network on n nodes. For $k \geqq 2$, there is an obvious lower bound on the number needed, namely $kn/2$ (for even kn) or $kn+1/2$ (for odd kn). These bounds can always be achieved.

This research was sponsored by the United States Air Force under Project RAND—Contract No. AF49 (638)-700.

References

1. R. G. Busacker and P. J. Gowen, A procedure for determining a family of minimal cost network flow patterns, O.R.O. Technical Paper, 15 (1961).

2. T. F. Cartaino and S. E. Dreyfus, Application of dynamic programming to the airplane minimum time-to-climb problem, Aero. Engr. Rev., 16 (1957) 74–77.

3. G. B. Dantzig, Application of the simplex method to a transportation problem, *Activity Analysis of Production and Allocation*, Cowles Commission Monograph 13, Wiley, New York, (1951) 359–373.

4. G. B. Dantzig and D. R. Fulkerson, Minimizing the number of tankers to meet a fixed schedule, Naval Res. Logist. Quart., 1 (1954) 217–222.

5. ———, On the max-flow min-cut theorem of networks, *Linear Inequalities and Related Systems*, Ann. of Math. Study 38, Princeton Univ. Press, (1956) 215–221.

6. G. B. Dantzig, D. R. Fulkerson, and S. Johnson, Solution of a large scale traveling salesman problem, Op. Res. 2, 393–410.

7. C. Derman and M. Klein, A note on the optimal depletion of inventory, Management Sci., 5 (1959) 210–214.

8. R. P. Dilworth, A decomposition theorem for partially ordered sets, Ann. of Math., 51 (1950) 161–166.

9. S. E. Dreyfus, A generalized equipment replacement study, J. Soc. Indust. Appl. Math., 8 (1960) 425–435.

10. L. R. Ford, Jr., Network flow theory, The RAND Corp., P-923 (1956).

11. L. R. Ford, Jr. and D. R. Fulkerson, Maximal flow through a network, Canad. J. Math., 8 (1956) 399–404.

12. ———, A simple algorithm for finding maximal network flows and an application to the Hitchcock problem, Canad. J. Math., 9 (1957) 210–218.

13. ———, A suggested computation for maximal multi-commodity network flows, Management Sci., 5 (1958) 97–101.

14. ———, Constructing maximal dynamic flows from static flows, Op. Res., 6 (1958) 419–433.

15. ———, *Flows in networks*, Princeton Univ. Press, 1962, 194 p.

16. D. R. Fulkerson, Increasing the capacity of a network: The parametric budget problem, Management Sci., 5 (1959) 472–483.

17. ———, An out-of-kilter method for minimal cost flow problems, J. Soc. Indust. Appl. Math., 9 (1961) 18–27.

18. ———, A network flow computation for project cost curves, Management Sci., 7 (1961) 167–178.

19. ———, A network flow feasibility theorem and combinatorial applications, Canad. J. Math., 11 (1959) 440–451.

20. D. R. Fulkerson and L. S. Shapley, Minimal k-arc-connected graphs, The RAND Corp., P-2371 (1961) 11 p.

21. D. Gale, A theorem on flows in networks, Pac. J. Math., 7 (1957) 1073–1082.

22. R. E. Gomory and T. C. Hu, Multi-terminal network flows, J. Soc. Indust. Appl. Math., 9 (1961) 551–571.

23. W. Gorham, An application of a network flow model to personnel planning, The RAND Corp., RM-2587 (1960) 85 p.

24. P. Hall, On representatives of subsets, J. Lond. Math. Soc., 10 (1935) 26–30.

25. F. L. Hitchcock, The distribution of a product from several sources to numerous localities, J. Math. and Phys., 20 (1941) 224–230.

26. A. J. Hoffman, Some recent applications of the theory of linear inequalities to extremal combinatorial analysis, Proc. Symposia Applied Math., 10 (1960).

27. T. C. Hu, Parallel sequencing and assembly line problems, Op. Res., 9 (1961) 841–849.

28. ———, The maximum capacity route problem, Op. Res., 9 (1961) 898–900.

29. W. S. Jewell, Optimal flow through networks with gains, Proc. Second International Conf. on Oper. Res., Aix-en-Provence, France 1960.

30. S. M. Johnson, Sequential production planning over time at minimum cost, Man. Sci., 3 (1957) 435–437.

31. ———, Optimal two- and three-stage production schedules with setup times included, Naval Res. Log. Q., 1 (1954) 61–68.

32. L. Kantorovitch and M. K. Gavurin, The application of mathematical methods in problems of freight flow analysis, *Collection of Problems Concerned with Increasing the Effectiveness of Transports*, Publ. of the Akad. Nauk SSSR, Moscow-Leningrad (1949) 110–138.

33. J. E. Kelley, Critical path planning and scheduling: mathematical basis, Op. Res., 9 (1961) 296–321.

34. T. C. Koopmans and S. Reiter, A model of transportation, *Activity Analysis of Production and Allocation*, Cowles Commission Monograph 13, Wiley, New York, 1951, 229–259.

35. D. König, *Theorie der endlichen und unendlichen Graphen*, Chelsea, 1950, 258 p.

36. H. W. Kuhn, The Hungarian method for the assignment problem, Naval Res. Log. Q., 2 (1955) 83–97.

37. J. B. Kruskal, Jr., On the shortest spanning subtree of a graph and the traveling salesman problem, Proc. Amer. Math. Soc., 7 (1956) 48–50.

38. G. J. Minty, Monotone networks, Proc. Roy. Soc., A, 257 (1960) 194–212.

39. A. Orden, The transshipment problem, Man. Sci., 3 (1956) 276–285.

40. M. Pollack, The maximum capacity route through a network, Op. Res., 8 (1960) 733–736.

41. A. F. Veinott, Jr. and H. M. Wagner, Optimal capacity scheduling—I and II, Op. Res., 10 (1962) 518–547.

42. D. F. Votaw, Jr. and A. Orden, The personnel assignment problem, Project SCOOP, Manual 10 (1952) 155–163.

43. H. M. Wagner and T. M. Whitin, Dynamic version of the economic lot-size model, Man. Sci., 5 (1958) 89–96.

QUADRATIC CONGRUENCES WITH AN ODD NUMBER OF SUMMANDS

ECKFORD COHEN, University of Tennessee

1. Introduction. It is well known [6, Chapters 6–11] that the formulas for the number of representations of an integer as a sum of an odd number of squares are, in general, more complicated than in the case of an even number. While this principle remains valid when the representations are taken with respect to a modulus, there are several important special cases of quadratic congruences with an odd number of unknowns for which the number of solutions is expressible by a quite simple formula. It is the purpose of this paper to exhibit some of these cases.

Let n be an integer and r a positive integer. Also, let $a_1, \cdots, a_{2m+1}$ be integers prime to r for $m \geqq 0$, and with $s=2m+1$, place

$$\alpha = (-1)^m a_1 \cdots a_s, \qquad (\alpha, r) = 1. \tag{1}$$

Denote by $B_m(n, r)$ the number of distinct solutions (mod r) of the congruence,

$$n \equiv a_1x_1^2 + \cdots + a_{2m+1}x_{2m+1}^2 \pmod r, \tag{2}$$

the a_i being subject to the condition in (1). We shall use Q to represent the set of square-free integers and $q(n)$ the characteristic function of Q, defined by $q(n)=1$ if $n \in Q$ and 0 if $n \notin Q$. The Jacobi symbol (mod r) will be denoted $\psi_r(n)$.

A particularly simple formula for $B_m(n, r)$ occurs when r is odd and n is relatively prime to r. In fact, we have the following elegant formula due, at least in an implicit form, to Minkowski:

$$B_m(n, r) = r^{2m} \sum_{d \mid r} \psi_d(\alpha n) \frac{q(d)}{d^m}, \quad (n, r) = 1, \ r \text{ odd}. \tag{3}$$

For an account of Minkowski's work on quadratic congruences, we refer to Bachmann [1, Chapter 7] and in particular to Sections 2 and 7 of that reference. Regarding the case $s=1$ of (3), in an equivalent form, we mention Landau [7, Theorems 71 and 87]. The history of the earlier development of the theory of quadratic congruences may be found in Dickson [6, Chapter 10].

The present paper is motivated by a belief that results such as the one contained in (3) should be more widely known. Although it is mainly expository, being based on an earlier paper [3], some of the details of proof are simplified by virtue of the special cases under consideration. In addition, several of the results of [3] are stated here in a simpler or more instructive form. We draw particular attention to the statement of Corollary 6 and its reformulation in Theorem 2 of Section 5.

As in [3], we use a method based on elementary properties of Gaussian sums and related sums. Preliminary material is collected in Section 2, the main results are proved in Section 3, special cases are discussed in Section 4, and a connection with totients is established in Section 5.

2. Preliminaries. We need some terminology and some further notation. A divisor d of r will be called *unitary* if d is relatively prime to the conjugate divisor r/d. The set of those integers which have no unitary square divisor other than 1 will be denoted by Q^*. The set Q^* is the "unitary" analogue of Q and consists of those integers whose canonical prime-power exponents are all odd (1, 2, 3, 5, 6, 7, 8, 10, 13, 15, $\cdots$). The *core* $\gamma(r)$ of r is the maximal divisor of r contained in Q; the maximal unitary divisor of r contained in Q is denoted $\gamma^*(r)$. Finally, $(n, r)_*$ will denote the largest unitary divisor of r which divides n, while (n,r) represents, as usual, the greatest common divisor of n and r.

The Kluyver function $\kappa(n, r)$ may be defined by the implicit relation,

$$\sum_{d|r} \kappa(n, d) = \eta(n, r) \overset{\text{def}}{=} \begin{cases} r & \text{if } r \mid n, \\ 0 & \text{if } r \nmid n. \end{cases} \tag{4}$$

The function $\kappa(n, r)$ is periodic (mod r) and has the trigonometric representation,

$$\kappa(n, r) = c(n, r) \overset{\text{def}}{=} \sum_{(x,r)=1} e(nx, r), \tag{5}$$

where $e(a, r) = \exp(2\pi i a/r)$, the summation ranging over a reduced residue system (mod r). The sum $c(n, r)$ is the familiar Ramanujan sum. Letting $\phi(r)$ denote the Euler totient and $\mu(r)$ the Möbius function, we recall that

$$\mu(r) = c(1, r), \qquad \phi(r) = c(0, r), \qquad \mu^2(r) = q(r). \tag{6}$$

For the explicit arithmetical representation of $\kappa(n, r)$ and a further discussion, we mention [4]; for our purposes it will suffice to have only the additional special result,

$$\kappa(n, r^2) = r\mu(r) \quad \text{if} \quad (n, r) = \gamma(r). \tag{7}$$

REMARK. In the rest of the paper r will be assumed *odd*.

The Gauss sum $G(n, r)$ is defined by

$$G(n, r) = \sum_{x \pmod r} e(nx^2, r), \tag{8}$$

and the associated sum $G^*(n, r)$ by

$$G^*(n, r) = \sum_{(x,r)=1} \psi_r(x)e(nx, r). \tag{9}$$

We note that $\psi_r(n)=0$ if $(n, r)\neq 1$, that $\psi_1(n)=1$ for all n, and that $\psi_r(n)=1$ if r is a square and $(n, r)=1$.

The following elementary properties of $G(n, r)$ will be needed (cf. [3, Section 2] and Walfisz [8, Section 1]):

$$G(n, r) = \psi_r(n)G(1, r) \quad \text{if} \quad (n, r) = 1, \tag{10}$$

$$G(nr', rr') = r'G(n, r), \tag{11}$$

$$G^2(1, r) = \psi_r(-1)r, \; G(1, r^2) = r. \tag{12}$$

We shall also require the following special formulas for $G^*(n, r)$. Let M denote the set of integral squares.

LEMMA 1. *If $\gamma(r)\,|\,n$, $(n, r)/(n, r)_*$ is in Q^*, and $r_1|\,r$, then*

$$G^*(n, r_1) = \begin{cases} c(n, r_1) & \text{if } r_1 \in M, \\ 0 & \text{otherwise.} \end{cases}$$

LEMMA 2. *If $(n, r)\,|\,\gamma^*(r)$, then*

$$G^*(n, r) = \begin{cases} G(n, r) & \text{if } r \in Q,\ (n, r) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

These results are special cases of a general relation for $G^*(n, r)$ proved in [3, Theorem 3]. We remark that in [3], $G(n, r)$ and $G^*(n, r)$ are designated $S(n, r)$ and $J(n, r)$, respectively.

3. The principal results. We now prove our main theorem. Recalling our previous notation and the Remark of Section 2, we have

THEOREM 1. *If $\gamma(r)\,|\,n$ and $(n, r)/(n, r)_*$ is contained in Q^*, then*

$$B_m(n, r) = r^{2m} \sum_{d^2 | r} \frac{\kappa(n, d^2)}{d^s}; \tag{13}$$

if $(n, r)\,|\,\gamma^*(r)$, then

$$B_m(n, r) = r^{2m} \sum_{d | r} \psi_d(\alpha n) \frac{\mu^2(d)}{d^m}. \tag{14}$$

Proof. Our proof is based on the fact that $B_m(n, r)$ is a periodic function of n with period r and hence has a finite Fourier expansion. In particular, by [1, Section 5], we may write

$$B_m(n, r) = \sum_{z \pmod r} a_r(z)e(zn, r), \tag{15}$$

where

$$a_r(z) = (1/r) \sum_{u \pmod r} e(-zu, r) B_m(u, r)$$
$$= (1/r) \sum_{x_i \pmod r,\ 1 \le i \le s} e(-z(a_1 x_1^2 + \cdots + a_s x_s^2), r),$$

so that by the definition of G,

$$a_r(z) = (1/r) \prod_{i=1}^{s} G(-a_i z, r). \tag{16}$$

In the summation (15), each $z \pmod r$ has a unique representation of the form $z = xr/d$, $x \pmod d$, $(x, d) = 1$, $d \mid r$. With such a representation, (16) becomes, on the basis of (10) and (11) and the oddness of s,

$$a_r(z) = (1/r)(r/d)^s \psi_d(-x a_1 \cdots a_s) G^s(1, d),$$

and hence by (15),

$$B_m(n, r) = r^{s-1} \sum_{d \mid r} \psi_d(-a_1 \cdots a_s) \left(\frac{G(1, d)}{d}\right)^s G^*(n, d).$$

Application of (12_1) leads then to

$$B_m(n, r) = r^{s-1} \sum_{d \mid r} \psi_d(-\alpha) \frac{G(1, d) G^*(n, d)}{d^{m+1}}. \tag{17}$$

This result holds for arbitrary n. We now specialize to the cases of the theorem.

In the first case, with $(n, r)/(n, r)_*$ assumed to be in Q^* and $\gamma(r)$ to divide n, we have by Lemma 1 and (17),

$$B_m(n, r) = r^{s-1} \sum_{d \mid r,\ d \in M} \frac{G(1, d) c(n, d)}{d^{m+1}},$$

and hence by (12_2), in this case,

$$B_m(n, r) = r^{s-1} \sum_{d^2 \mid r} \frac{c(n, d^2)}{d^{2m+1}}.$$

The first part of the theorem results now by (5).

In the second case, in which (n, r) is assumed to be a factor of $\gamma^*(r)$, we have by (17), Lemma 2, and (10),

$$B_m(n, r) = r^{s-1} \sum_{d \mid r, d \in Q, (d,n)=1} \psi_d(-\alpha n) \frac{G^2(1, d)}{d^{m+1}},$$

and therefore by (12_1),

$$B_m(n, r) = r^{s-1} \sum_{d \mid r, (d,n)=1} \psi_d(\alpha n) \frac{q(d)}{d^m}.$$

The relative primality condition is superfluous and hence (14) is proved.

4. Special cases. It is noted that the condition of the second case of the theorem is satisfied if $(n, r)=1$; hence the result stated in the Introduction. The condition of (14) is also satisfied for a square-free modulus r:

COROLLARY 1. *If* $r \in Q$, *then for all* n, $B_m(n, r)=r^{2m}\sum_{d|r}\psi_d(\alpha n)/d^m$.

If $r|n$, the latter sum is 0 unless $d=1$; thus

COROLLARY 2. *If* $r \in Q$ *and* $n \equiv 0 \pmod r$, *then* $B_m(n, r)=r^{2m}$.

It is also of interest to record the case $s=1$ of (14).

COROLLARY 3 ($m=0$). *If* $(n, r)|\gamma^*(r)$, *then*

$$B_0(n, r) = \sum_{d|r} \psi_d(\alpha n)\mu^2(d).$$

The first case of the theorem arises when $r|n$. In this case (13) yields, by virtue of (5), (6), and the fact that $\phi(r^2)=r\phi(r)$,

COROLLARY 4. *If* $n \equiv 0 \pmod r$, *then* $B_m(n, r)=r^{2m}\sum_{d^2|r}\phi(d)/d^{2m}$.

As for the case $s=1$, we have

COROLLARY 5 ($m=0$). *If* $n \equiv 0 \pmod r$, *then* $B_0(n, r)=\sum_{d^2|r}\phi(d)$.

The condition of the first case of the theorem is also satisfied when $\gamma(r)\|n$. More generally, by (7) and (13) it follows that

COROLLARY 6. *If* $(n, r)=\gamma(r)$, *then*

$$B_m(n, r) = r^{2m}\sum_{d^2|r}\frac{\mu(d)}{d^{2m}}.$$

Note that Corollary 2 is a consequence of either Corollary 4 or 6.

5. A connection with totient functions. In case $m=0$, Corollary 6 reduces to the familiar relation,

$$q(r) = \sum_{d^2|r}\mu(d). \tag{18}$$

We use this formula in establishing a combinatorial evaluation of $B_m(n, r)$ in the case of Corollary 6.

Let us define $\phi_t(r)$ for arbitrary positive integers t to be the number of t-tuples of integers $b_1, \cdots, b_t$, distinct (mod r), such that the greatest common divisor $(b_1, \cdots, b_t, r)$ of $b_1, \cdots, b_t$ and r is contained in Q.

We now prove

THEOREM 2. *If* $(n, r)=\gamma(r)$, *then for* $s>1$,

$$B_m(n, r) = r^m\phi_m(r). \tag{19}$$

Proof. By Corollary 6 of Section 4,

$$\frac{B_m(n,r)}{r^m} = \sum_{d^2|r}\mu(d)(r/d^2)^m = \sum_{d^2|r}\mu(d)\left(\sum_{x_i \pmod{r/d^2},1\le i\le m} 1\right)$$

$$= \sum_{d^2|r}\mu(d)\left(\sum_{b_i \pmod{r},d^2|b_i,1\le i\le m} 1\right) = \sum_{b_i \pmod{r},1\le i\le m}\ \sum_{d^2|b_i,d^2|r}\mu(d).$$

Therefore by (18),

$$\frac{B_m(n,r)}{r^m} = \sum_{b_i \pmod{r},1\le i\le m} q((b_1,\cdots,b_m,r)) = \sum_{(b_1,\cdots,b_m,r)\in Q,b_i \pmod{r}} 1,$$

which is $\phi_m(r)$. Place $\phi(r)=\phi_1(r)$.

COROLLARY ($m=1$). *If* $(n,r)=\gamma(r)$, *then* $B_1(n,r)=r\phi(r)$.

For an application of the function $\phi(r)$, we mention Cohen and Robinson [5].

The author is indebted to the National Science Foundation for its support in the completion of this paper.

References

1. Paul Bachmann, Zahlentheorie, vol. 4: Arithmetik der quadratischen Formen, Part 1, Leipzig, 1898.
2. Eckford Cohen, Rings of arithmetic functions, Duke Math. J., 19 (1952) 115–129.
3. ———, Rings of arithmetic functions, II, The number of solutions of quadratic congruences, Duke Math. J., 21 (1954) 9–28.
4. ———, The elementary arithmetical functions, Scripta Mathematica, 25 (1960) 221–227.
5. Eckford Cohen and Richard L. Robinson, Square-free integers in residue classes, to appear.
6. Leonard Eugene Dickson, History of the Theory of Numbers, vol. II, repr. by Chelsea, 1952.
7. Edmund Landau, Vorlesungen über Zahlentheorie, vol. I, repr. by Chelsea, 1950.
8. Arnold Walfisz, On the representation of numbers by sums of squares, Transl. of the Amer. Math. Soc., Series 2, vol. 3, pp. 163–248.

ON SOME SUMS AND INTEGRALS INVOLVING BESSEL FUNCTIONS

ELDON HANSEN, Lockheed, Missiles and Space Div.

Introduction. In problem 5099 of this MONTHLY, Mitrinovic [1] posed the problem of summing the series

$$\sum_{k=0}^{\infty}(2k+p)J_k(x)J_{p+k}(x), \tag{1}$$

when p is an even integer. The case p an odd integer is known (see the editorial comment on page 449 of [2]).

In this paper we consider the more general sum

$$S_m(p, q, a, x, y, n) = \sum_{k=1}^{n} k^m a^k J_{p+k}(x) J_{q+k}(y), \tag{2}$$

where n may be finite or infinite and where m is a nonnegative integer. The parameters p, q, x, y, and a are, in general, arbitrary real or complex numbers. For simplicity, we shall assume x, y and a are nonzero.

We shall evaluate the sum in equation (2) and some similar sums in three particular cases. We conclude by evaluating certain integrals involving Bessel functions which have been expressed in terms of these sums.

To conserve space, we shall write only as many of the (leading) arguments of S_m as are required for clarity. Arguments not shown will be assumed to be the same as in equation (2).

Our main results will be derived using only the recursion relation

$$2rJ_r(z) = z[J_{r+1}(z) + J_{r-1}(z)] \tag{3}$$

and hence hold equally well for other functions satisfying this relation.

Some basic relations. Using equation (3) with $r = k + p$,

$$\begin{aligned} S_m + pS_{m-1} &= \sum_{k=1}^{n} k^{m-1} a^k (k + p) J_{p+k}(x) J_{q+k}(y) \\ &= \frac{x}{2} \sum_{k=1}^{n} k^{m-1} a^k [J_{p+k+1}(x) + J_{p+k-1}(x)] J_{q+k}(y) \\ &= \frac{x}{2} [S_{m-1}(p + 1) + S_{m-1}(p - 1)]. \end{aligned} \tag{4}$$

Similarly,

$$S_m + qS_{m-1} = \frac{y}{2} [S_{m-1}(p, q + 1) + S_{m-1}(p, q - 1)]. \tag{5}$$

Note that equation (4) or (5) relates S_m to corresponding sums with m replaced by $m-1$. Hence if we can find S_0 or S_1 for appropriate values of p and q, we can recursively find S_m for any positive integer m. We now proceed to evaluate S_0 or S_1.

From the defining equation (2),

$$\begin{aligned} aS_0(p, q + 1) &= \sum_{k=1}^{n} a^{k+1} J_{p+k}(x) J_{q+k+1}(y) \\ &= \sum_{k=2}^{n+1} a^k J_{p+k-1}(x) J_{q+k}(y) \\ &= S_0(p - 1, q) + A(p, q, a, x, y, n), \end{aligned} \tag{6}$$

where we have used the abbreviation

$$A(p,\, q,\, a,\, x,\, y,\, n) = a^{n+1}J_{p+n}(x)J_{q+n+1}(y) - aJ_p(x)J_{q+1}(y). \tag{7}$$

As in the case of S_m, we shall write only those arguments of A which are necessary for clarity. For later use, we rewrite equation (6) with p and q replaced by $p+1$ and $q-1$, respectively. Then

$$aS_0(p+1) = S_0(p,\, q-1) + A(p+1,\, q-1). \tag{8}$$

Using equations (4) and (5) with $m=1$, combined with equations (6) and (8), we have four equations in the six unknowns S_0, $S_0(p+1)$, $S_0(p-1)$, $S_0(p,\, q+1)$, $S_0(p,\, q-1)$ and S_1. For special values of some of the parameters p, q, a, x and y, we can isolate one of these six unknowns. These cases yield the following results which are easily verified:

$$\begin{aligned} S_0(p,\, q,\, \pm 1,\, x,\, \pm x) &= \sum_{k=1}^{n} (\pm 1)^k J_{p+k}(x)J_{q+k}(\pm x) \\ &= \frac{x}{2(q-p)}\,[A(p,\, q,\, \pm 1,\, x,\, \pm x) \\ &\quad \mp A(p+1,\, q-1,\, \pm 1,\, x,\, \pm x)] \qquad (p \neq q), \end{aligned} \tag{9}$$

$$\begin{aligned} S_0(p+1,\, p,\, a,\, x,\, ax) &= \sum_{k=1}^{n} a^k J_{p+k+1}(x)J_{p+k}(ax) \\ &= \frac{1}{a^2-1}\,[aA(p+1,\, p-1,\, a,\, x,\, ax) \\ &\quad - A(p,\, p,\, a,\, x,\, ax)] \qquad (a^2 \neq 1), \end{aligned} \tag{10}$$

$$\begin{aligned} S_1(p,\, q,\, \pm 1,\, x,\, \pm qx/p) &= \sum_{k=1}^{n} (\pm 1)^k k J_{p+k}(x)J_{q+k}(\pm qx/p) \\ &= \frac{qx}{2(p-q)}\,[A(p,\, q,\, \pm 1,\, x,\, \pm qx/p) \\ &\quad \mp A(p+1,\, q-1,\, \pm 1,\, x,\, \pm qx/p)] \qquad (p \neq q). \end{aligned} \tag{11}$$

The two equations expressed by (9) by using the upper or the lower signs are identical. This can be shown using the relation $J_r(e^{m\pi i}z) = e^{mr\pi i}J_r(z)$. The same is true of (11). But if some other function $F_r(z)$, satisfying equation (3), were used instead of $J_r(z)$, this may not be the case. Hence we have included both signs for completeness. For simplicity, however, we shall use only the upper signs in what follows.

Now $nc^nJ_{n+r}(cx) \to 0$ as $n \to \infty$ for arbitrary (but fixed) r, x and c. Hence the terms of the sums in equations (9), (10) and (11) approach zero as $k \to \infty$. Therefore, also, each right hand member in equations (9), (10) and (11) is uniformly bounded for large n. We thus conclude that each of these series converges as

$n \to \infty$ and we obtain, respectively,

$$\sum_{k=1}^{\infty} J_{p+k}(x)J_{q+k}(x) = \frac{x}{2(q-p)}\left[J_{p+1}(x)J_q(x) - J_p(x)J_{q+1}(x)\right], \quad (p \neq q), \tag{12}$$

$$\sum_{k=1}^{\infty} a^k J_{p+k+1}(x)J_{p+k}(ax) = \frac{a}{a^2-1}\left[J_p(x)J_{p+1}(ax) - aJ_{p+1}(x)J_p(ax)\right], (a^2 \neq 1), \tag{13}$$

and

$$\sum_{k=1}^{\infty} kJ_{p+k}(x)J_{q+k}(qx/p) = \frac{qx}{2(p-q)}\left[J_{p+1}(x)J_q(qx/p) - J_p(x)J_{q+1}(qx/p)\right], \quad (p \neq q). \tag{14}$$

Mixed cases. So far, we have used equations (4), (5), (6) and (8) only to solve for one of the sums S_0, $S_0(p+1)$, $S_0(p-1)$, $S_0(p, q+1)$, $S_0(p, q-1)$, or S_1 alone. If we solve for a combination of two or more of these sums, we can obtain results with less restriction on the parameters. For example, we can get

$$\begin{aligned}(x-y)S_1(p, q, 1) &+ (qx - py)S_0(p, q, 1) \\ &= \sum_{k=1}^{n}\left[(x-y)k + (qx-py)\right]J_{k+p}(x)J_{k+q}(y) \\ &= \frac{xy}{2}\left[A(p, q, 1) - A(p+1, q-1, 1)\right]\end{aligned} \tag{15}$$

with no restrictions on p, q, x or y.

Other sums. Having considered S_m, it seems natural to consider the similar sums

$$T_m(p, q, a, x, y, n) = \sum_{k=1}^{n} k^m a^k J_{p-k}(x)J_{q+k}(y) \tag{16}$$

and

$$U_m(p, q, a, x, y, n) = \sum_{k=1}^{n} k^m a^k J_{p-k}(x)J_{q-k}(y). \tag{17}$$

Let us first consider T_m.

Corresponding to equations (4), (5) and (6) we obtain

$$T_m = pT_{m-1} - \frac{x}{2}\left[T_{m-1}(p+1) + T_{m-1}(p-1)\right], \tag{18}$$

$$T_m = -qT_{m-1} + \frac{y}{2}\left[T_{m-1}(p, q+1) + T_{m-1}(p, q-1)\right], \tag{19}$$

and

$$aT_0(p,\ q+1) = T_0(p+1) + B, \tag{20}$$

where

$$B(p,\ q,\ a,\ x,\ y) = a^{n+1}J_{p-n}(x)J_{q+n+1}(y) - aJ_p(x)J_{q+1}(y). \tag{21}$$

Corresponding to equations (9), (10) and (11), we obtain from equations (18), (19) and (20),

$$\begin{aligned} T_0(p,\ q,\ 1,\ x,\ -x) &= (-1)^q \sum_{k=1}^{n} (-1)^k J_{p-k}(x)J_{q+k}(x) \\ &= \frac{x}{2(p+q)}\,[B(p-1,\ q-1,\ 1,\ x,\ -x) \\ &\quad - B(p,\ q,\ 1,\ x,\ -x)] \qquad (q \neq -p), \end{aligned} \tag{22}$$

$$\begin{aligned} T_0(p-1,\ -p,\ a,\ x,\ -ax) &= \sum_{k=1}^{n} a^k J_{p-k-1}(x)J_{k-p}(-ax) \\ &= \frac{1}{1-a^2}\,[B(p,\ -p,\ a,\ x,\ -ax) \\ &\quad - aB(p-1,\ -p-1,\ a,\ x,\ -ax)] \qquad (a^2 \neq 1), \end{aligned} \tag{23}$$

and

$$\begin{aligned} T_1(p,\ q,\ 1,\ x,\ qx/p) &= \sum_{k=1}^{n} kJ_{p-k}(x)J_{q+k}(qx/p) \\ &= \frac{qx}{2(p+q)}\,[B(p,\ q,\ 1,\ x,\ qx/p) \\ &\quad - B(p-1,\ q-1,\ 1,\ x,\ qx/p)] \qquad (q \neq -p). \end{aligned} \tag{24}$$

To obtain equation (22), we used the fact that $J_r(-x) = (-1)^r J_r(x)$.

Similarly,

$$U_m = pU_{m-1} - \frac{x}{2}\,[U_{m-1}(p+1) + U_{m-1}(p-1)], \tag{25}$$

$$U_m = qU_{m-1} - \frac{y}{2}\,[U_{m-1}(p,\ q+1) + U_{m-1}(p,\ q-1)], \tag{26}$$

and

$$U_0(p,\ q+1) = aU_0(p-1) + C, \tag{27}$$

where

$$C(p,\ q,\ a,\ x,\ y) = aJ_{p-1}(x)J_q(y) - a^{n+1}J_{p-n-1}(x)J_{q-n}(y). \tag{28}$$

From equations (25), (26) and (27),

$$
\begin{aligned}
U_0(p, q, 1, x, x) &= \sum_{k=1}^{n} J_{p-k}(x)J_{q-k}(x) \\
&= \frac{x}{2(q-p)}\,[C(p, q, 1, x, x) \\
&\quad - C(p+1, q-1, 1, x, x)] \qquad (q \neq p),
\end{aligned} \tag{29}
$$

$$
\begin{aligned}
U_0(p-1, p, a, x, ax) &= \sum_{k=1}^{n} a^k J_{p-k-1}(x)J_{p-k}(ax) \\
&= \frac{1}{1-a^2}\,[aC(p, p, a, x, ax) \\
&\quad - C(p+1, p-1, a, x, ax)] \qquad (a^2 \neq 1),
\end{aligned} \tag{30}
$$

and

$$
\begin{aligned}
U_1(p, q, 1, x, qx/p) &= \sum_{k=1}^{n} kJ_{p-k}(x)J_{q-k}(qx/p) \\
&= \frac{qx}{2(q-p)}\,[C(p, q, 1, x, qx/p) \\
&\quad - C(p+1, q-1, 1, x, qx/p)] \qquad (q \neq p).
\end{aligned} \tag{31}
$$

Limiting cases. In equation (10), (13), (23) and (30) we had the restriction $a^2 \neq 1$. This excluded case is, in each of these equations, a special case of the equation immediately preceding it. This is obvious for $a=1$. For $a=-1$, it follows easily by using $J_r(-z)=e^{2\pi i r}J_r(z)$.

In other equations, we had the restriction $p \neq q$ or $p \neq -q$. For each of these equations, we can obtain the excluded case. We illustrate this by considering equations (12) and (14). In equation (12), let $q \to p$ and use L'Hospital's rule. We get

$$
\sum_{k=1}^{\infty} [J_{p+k}(x)]^2 = \frac{x}{2}\left[J_{p+1}(x)\,\frac{\partial J_p(x)}{\partial p} - J_p(x)\,\frac{\partial J_{p+1}(x)}{\partial p}\right]. \tag{32}
$$

For $\mathrm{Re}(p) > 0$, we can rewrite this as

$$
\sum_{k=1}^{\infty} [J_{p+k}(x)]^2 = p\int_0^x [J_p(t)]^2\,\frac{dt}{t} - \frac{1}{2}\,[J_p(x)]^2 \tag{33}
$$

which is given as equation (5), section 5.51 of [**4**]. It is known that the right members of (32) and (33) are equal. See equation (6), page 255 of [**6**].

If we let $q \to p$ in equation (14), the series becomes $S_1(p, p, 1, x, x, \infty)$ which

can be obtained by substituting equation (32) and appropriate cases of equation (12) into equation (4) (with $m=1$).

Known special cases. We have pointed out that a result equivalent to equation (32) was known provided $\mathrm{Re}(p)>0$. We now indicate other known special cases of equations derived above.

Equation (20), section 7.15 of [3] is:

$$\sum_{k=1}^{\infty} J_{k+p}(x)J_{k-p}(x) = \frac{\sin \pi p}{2\pi p} - \frac{1}{2} J_p(x)J_{-p}(x). \tag{34}$$

This is the special case $q=-p$ of equation (12) as can be shown using the relation

$$2\pi p J_p(x)J_{-p}(x) + \pi x[J_p(x)J_{1-p}(x) - J_{p+1}(x)J_{-p}(x)] = 2 \sin \pi p.$$

If we let $p \to 0$ in equation (34), we obtain the well-known result

$$\sum_{k=1}^{\infty} [J_k(x)]^2 = \frac{1}{2} - \frac{1}{2}[J_0(x)]^2. \tag{35}$$

Equation (21), section 7.15 of [3] is:

$$\sum_{k=1}^{\infty} J_{p-k}(x)J_{p+k}(x) \cos kt = \frac{1}{2} J_{2p}[2x \cos (t/2)] - \frac{1}{2}[J_p(x)]^2, \tag{36}$$

which is valid for $\mathrm{Re}(p)>0$ and $-\pi \leqq t \leqq \pi$. Letting $t=\pi$ yields

$$\sum_{k=1}^{\infty} (-1)^k J_{p-k}(x)J_{p+k}(x) = -\frac{1}{2}[J_p(x)]^2.$$

This equation also follows from equation (22).

The series $S_m(0, 0, 1, x, x)=\sum_{k=1}^{\infty} k^m[J_k(x)]^2$ are known for m even. They follow from equations (3), page 37 of [4] and are all polynomials in x. For m odd these series are apparently not known. They follow recursively however, by use of equations (4), (12) and (35). In these cases, they are not polynomials. For example,

$$\sum_{k=1}^{\infty} k[J_k(x)]^2 = \frac{x^2}{2}\{[J_0(x)]^2 + [J_1(x)]^2\} - \frac{x}{2} J_0(x)J_1(x).$$

Evaluation of certain integrals. Certain integrals involving Bessel functions have been evaluated in terms of some of the series discussed above. By summing these series, we have thus evaluated these integrals in more simple form. We give two illustrative results without derivation. They follow easily from previous equations.

Using equation (33), page 982 of [5] and our equation (15) with $p=1/4$, $q=3/4$ and $y=-x$, we obtain

$$\int_0^x J_{1/4}(t)J_{-1/4}(t)\cos 2(x-t)t^{1/2}dt$$

$$= x^{1/2}\sum_{k=0}^{\infty}(-1)^k(2k+1)J_{k+3/4}(x)J_{k+1/4}(x)$$

$$= x^{1/2}J_{3/4}(x)J_{1/4}(x) - \frac{1}{2}x^{3/2}[J_{1/4}(x)J_{7/4}(x) + J_{5/4}(x)J_{3/4}(x)].$$

Using equation (332), page 982 of [5], and our equation (15), we get

$$\int_0^x J_p(t)J_{p-1/2}(t)t^{1/2}dt = x^{1/2}\sum_{k=0}^{\infty}J_{p+k}(x)J_{p+k+1/2}(x)$$

$$= x^{1/2}J_p(x)J_{p+1/2}(x)+x^{3/2}[J_{p+1}(x)J_{p+1/2}(x)-J_p(x)J_{p+3/2}(x)].$$

This holds for $\mathrm{Re}(p) > -1/2$.

For other similar cases, see [6].

References

1. Problem 5099, Summation of products of Bessel functions, proposed by D. S. Mitrinovic, this MONTHLY, 70 (1963) 445.
2. Solution to problem 5099 by Eldon Hansen, this MONTHLY, 71 (1964) 448–449.
3. Erdelyi, et al., Higher transcendental functions, vol. 2, McGraw-Hill, New York, 1953.
4. G. N. Watson, A treatise on the theory of Bessel functions, 2nd ed. Cambridge University Press, 1958.
5. J. G. Rutgers, Nederl. Akad. Wetensch. Proc., 44 (1941) 978–988.
6. Yudell L. Luke, Integrals of Bessel functions, McGraw-Hill, New York, 1962.

GENERALIZED INVERTIBLE SPACES

DANG XUAN HONG, Peabody College and Vanderbilt University

1. Introduction. Doyle and Hocking [2] define a topological space X to be *invertible*, provided for each nonempty open subset U of X, there exists a homeomorphism h of X onto X such that $h(X-U)\subset U$. Several theorems in the literature are of the following type: if there is a nonempty open set U of X such that U has property P, then X has property P. The purpose of this paper is to generalize the notion of an invertible space and to extend to this more general setting many of the results in the literature.

DEFINITION. *A topological space X is a generalized$_1$ (generalized$_2$) invertible space provided there exists a proper (for each) open subset U of X and there is (there exists) a homeomorphism h of X onto X such that for each $x\in X$, $h^n(x)\in U$ for some integer $n(x)$. The pair (U, h) is called an inverting pair for X.*

Primarily we will confine our attention to generalized$_1$ invertible spaces. Each Euclidean space E^n is a generalized$_1$ invertible space, although E^n is not an

invertible space. In fact, Doyle and Hocking [3] have shown that the n-sphere is the only invertible n-manifold. In Section 2, we show that if (U, h) is an inverting pair for X, and if U has certain properties, then X also has these properties. However, U may have some properties (e.g. paracompactness, metrizability, etc.) which are not possessed by X, and in Section 3, we impose further restrictions on U so that some additional properties of U are carried over to X.

The author wishes to thank Professor B. F. Bryant of Vanderbilt University for his encouragement and advice during the preparation of this paper.

2. Throughout the remainder of the paper, we suppose that (U, h) is an inverting pair for X. This section is devoted to the proofs of the following theorems:

If U has property P, then X has property P, where P is

(1) T_0.
(2) T_1.
(3) *separable.*
(4) *first countable.*
(5) *second countable.*
(6) *Lindelöf.*
(7) *first category.*
(8) *Baire.*

Proofs. (1) Let $x, y \in X$ with $x \neq y$. If $x, y \in U$, then the conclusion holds. If $x \in U$ and $y \notin U$, then U is the required open set. (2) Let $x, y \in X$ with $x \neq y$. If $x, y \in U$, then the conclusion holds. Next, suppose $x \in U$ and $y \notin U$. There exists n such that $h^n(y) \in U$; if $h^n(x) \in U$, then there exists an open set W such that $h^n(y) \in W$ and $h^n(x) \notin W$. Hence $y \in h^{-n}(W)$ and $x \notin h^{-n}(W)$, and the conclusion holds. Finally, suppose that $x, y \notin U$, and $h^n(x) \in U$. If $h^n(y) \in U$, we apply the first part of the proof; if $h^n(y) \notin U$, we apply the second part of the proof.

(3) If A is a countable dense subset of U, then

$$B = \{h^n(a) \mid n \text{ an integer}, a \in A\}$$

is a countable dense subset of X.

(4) Let $x \in X$ and $h^n(x) \in U$. If $\{U_i\}$ is a countable basis about $h^n(x)$, then $\{h^{-n}(U_i)\}$ is a countable basis about x.

(5) If $\{U_i \mid i \text{ a positive integer}\}$ is a countable basis for U, then $K = \{h^j(U_i) \mid j$ an integer and i a positive integer$\}$ is a countable basis for X.

(6) Let $\{O_\alpha\}$ be an open cover of X. For each integer n, $\{h^n(O_\alpha) \cap U\}$ is an open cover of U. Hence there exists a countable subcover

$$\{h^n(O_{n_i}) \cap U \mid i = 1, 2, 3, \cdots\}.$$

The collection $\{O_{n_i} \mid n$ an integer and $i = 1, 2, 3, \cdots\}$ is a countable subcover of the cover $\{O_\alpha\}$.

(7) If a set A is nowhere dense in U, then A (and hence $h(A)$) is nowhere dense in X. Hence, if $U = \bigcup_{i=1}^\infty A_i$ where each A_i is nowhere dense in U, then $X = \bigcup h^n(A_i)$ for n an integer and $i = 1, 2, 3, \cdots$, and each $h^n(A_i)$ is nowhere dense in X. Thus X is of the first category.

(8) A topological space S is a Baire space if for each countable collection $\{O_i\}$, where each O_i is open and dense in S, $\bigcap_{i=1}^\infty O_i$ is dense in S [1]. We first

establish that if each point of X is contained in an open set which is a Baire space, then X is a Baire space. Let $\{O_i\}$ be a countable collection of open sets, each of which is dense in X. Let $x \in X$, and let V be an open Baire space containing x. Let W be an arbitrary open set containing x. Since each O_i is dense in X, each $O_i \cap V$ is dense in V. Since V is a Baire space, $\bigcap_{i=1}^{\infty} (O_i \cap V)$ is dense in V. Thus

$$W \cap \left(\bigcap_{i=1}^{\infty} O_i \right) \supseteq W \cap \left[\bigcup_{i=1}^{\infty} (O_i \cap V) \right] = (W \cap V) \cap \left[\bigcap_{i=1}^{\infty} (O_i \cap V) \right] \neq \varnothing.$$

Therefore $\bigcap_{i=1}^{\infty} O_i$ is dense in X, and hence X is a Baire space. Now for each $x \in X$, $h^n(x) \in U$ for some n; but U is a Baire space, and hence $x \in h^{-n}(U)$ which is an open Baire space. The first part of the proof implies that X is a Baire space.

REMARK. Let $A = \{a_i \mid i = 1, 2, 3, \cdots\}$, and let $X = \{a, b\} \cup A$. We define each subset of A to be open, and if $a \in W \subset X$ (or $b \in W \subset X$), then W is open if and only if W contains all but a finite number of the members of A. Then (U, h) is an inverting pair for X, where $U = \{a\} \cup A$, and

$$h(x) = \begin{cases} x & \text{if } x \in A, \\ a & \text{if } x = b, \\ b & \text{if } x = a. \end{cases}$$

U is T_2, T_3, paracompact, and metrizable, but X is not T_2, T_3, paracompact, nor metrizable. Thus, in the next section we place further restrictions on U so that whenever U has the above properties, so does X.

3. THEOREM 9. *Suppose $U \subset A$, where A is closed. If A is T_i $(i = 2, 3)$, then X is T_i $(i = 2, 3)$.*

Proof. We only give the proof for $i = 3$. Let V be an open set containing x, and let $h^n(x) \in U$. Hence $h^n(x) \in h^n(V) \cap U \subset U \subset A$; so there exists W (open in A and contained in U, and hence open in X) such that $h^n(x) \in W \subset \overline{W} \subset h^n(V) \cap U$ (since A is closed, the closure of W relative to A is the closure of W in X). Thus $x \in h^{-n}(\overline{W}) \subset h^{-n}(h^n(V) \cap U) \subset V$, so X is T_3.

THEOREM 10. *Let X be an invertible space. If X contains a nonempty open set U such that U is T_i $(i = 2, 3)$, then X is T_i $(i = 2, 3)$.*

This theorem is due to Doyle and Hocking [2] for T_2 and to Levine [7] for T_3.

THEOREM 11. *Suppose $U \subset A$, where A is closed. Then A is paracompact if and only if X is paracompact.*

Proof. One part of the theorem is the observation that a closed subset of a paracompact space is paracompact ([4], p. 78). Next, let $\{O_\alpha\}$ be an open cover of X. With Kelley, we require a paracompact space to be T_3. (We remind the reader that Kelley uses regular for what we have called T_3). Thus A is T_3, and

hence X is T_3 by Theorem 9. We shall show that $\{O_\alpha\}$ has an open σ-locally finite refinement which will establish that X is paracompact. For each n, $\{h^n(O_\alpha)\cap A\}$ is a relatively open cover of A. Thus there exists a relatively open locally finite refinement $\{W_{\beta_n}\cap A\}$ where each W_{β_n} is open in X. Thus for each fixed n, the collection $\{W_{\beta_n}\cap U\}$ is also locally finite relative to A and $W_{\beta_n}\cap U$ is open in X. Since A is closed, one can show that for each n, the family $\{h^{-n}(W_{\beta_n}\cap U)\}$ is locally finite relative to X. Also, for each β_n there exists α such that $W_{\beta_n}\cap U\subset h^n(O_\alpha)\cap A$; hence

$$h^{-n}(W_{\beta_n} \cap U) \subset h^{-n}(h^n(O_\alpha) \cap A) \subset O_\alpha.$$

Therefore $\bigcup\{h^{-n}(W_{\beta_n}\cap U) \mid n \text{ an integer}\}$ is a σ-locally finite refinement of $\{O_\alpha\}$, and hence X is paracompact.

This proof is also valid if X is a generalized$_2$ invertible space provided X is T_3. A similar statement holds concerning Theorem 12 below.

COROLLARY. *Let X be an invertible space. If X contains a nonempty open set U such that U is paracompact, then X is paracompact.*

Proof. U is paracompact, and hence T_3; thus X is T_3 by Theorem 10. Therefore there exists a nonempty open set V such that $V\subset\overline{V}\subset U$. But U is paracompact, so $\overline{V}$ is paracompact. Also there exists a homeomorphism h such that (V, h) is an inverting pair for X. Hence, by the above theorem, X is paracompact.

THEOREM 12. *Suppose $U\subset A$, where A is closed. If A is metrizable, then X is metrizable.*

Proof. A is regular (i.e. T_1 and T_3), and hence X is regular by Theorems 2 and 10. A regular space is metrizable if and only if it has a σ-locally finite basis ([6], p. 127). Hence, for each positive integer n, there exists a collection $\{W_{\beta_n}\cap A\}$ which is locally finite relative to A such that each W_{β_n} is open in X and $\bigcup\{W_{\beta_n}\cap A \mid n \text{ a positive integer}\}$ is a basis for A. Hence $\bigcup\{W_{\beta_n}\cap U \mid n \text{ a positive integer}\}$ is a basis for U, and for each n, the family $\{W_{\beta_n}\cap U\}$ is locally finite relative to A. Furthermore, the family

$$\bigcup \{h^m(W_{\beta_n} \cap U) \mid n \text{ a positive integer and } m \text{ an integer}\}$$

is a basis for X. Since A is closed, one can show that for m and n fixed, the family $\{h^m(W_{\beta_n}\cap U)\}$ is a locally finite family relative to X. Thus $\bigcup\{h^m(W_{\beta_n}\cap U) \mid n$ a positive integer and m an integer$\}$ is a σ-locally finite basis for X, and hence X is metrizable.

COROLLARY. (GRAY [5]) *Let X be an invertible space. If X contains a nonempty open set U such that U is metrizable, then X is metrizable.*

Proof. U is metrizable, so U is T_3; hence X is T_3 by Theorem 10. Therefore there exists a nonempty open set V such that $V\subset\overline{V}\subset U$, and hence $\overline{V}$ is metrizable. Also there exists a homeomorphism h such that (V, h) is an inverting pair

for X. Thus, by the above theorem, X is metrizable.

If X is the integers, U a single point, and h the shift homeomorphism, then (U, h) is an inverting pair for X and U is compact, but X is not compact. However, one has the following

THEOREM 13. *If $U \subset A$, where A is compact, and if there exists N such that for each $x \in X$, $h^n(x) \in U$ for some n with $|n| < N$, then X is compact.*

Proof. Let $\{O_\alpha\}$ be an open cover of X. For each n, $|n| < N$, $\{h^n(O_\alpha) \cap A\}$ is an open cover of A; thus there is a finite subcover

$$\{h^n(O_{n_i}) \cap A \mid i = 1, 2, \cdots, I_n\}.$$

The family $\{O_{n_i} \mid |n| < N, i = 1, 2, \cdots, I_n\}$ is a finite cover of X.

COROLLARY. (DOYLE AND HOCKING [2]) *If an invertible space X contains a nonempty open set U whose closure $\overline{U}$ is compact, then X is compact.*

References

1. N. Bourbaki, Topologie générale, Actualités Sci. Ind., 1045 (1948) Paris.
2. P. H. Doyle and J. G. Hocking, Invertible spaces, this MONTHLY, 68 (1961) 959–965.
3. ———, A characterization of the n-sphere, Mich. Math. J., 7 (1960) 199–200.
4. J. G. Hocking and G. S. Young, Topology, Addison-Wesley, Reading, Mass. 1961.
5. William J. Gray, On the metrizability of invertible spaces, this MONTHLY, 71 (1964) 533–534.
6. J. L. Kelley, General Topology, Van Nostrand, New York, 1955.
7. Norman Levine, Some remarks on invertible spaces, this MONTHLY, 70 (1963) 181–183.

ON THE NUMBER OF TOPOLOGIES ON A FINITE SET

V. KRISHNAMURTHY, University of Illinois, Urbana

1. Introduction. Let n be a finite positive integer. Let $N = \{1, 2, \cdots, n\}$ and $E = \{x_i, i \in N\}$. Let $f(n)$ be the number of distinct topologies that can be defined on E. Note that we shall not be taking homeomorphisms into account and so we shall be counting two homeomorphic topologies as distinct provided the open sets are not identical. So $f(1) = 1$ and $f(2) = 4$. Also, with some effort one can see that $f(3) = 29$. But $f(4)$, as far as we know, has not been computed. It is also clear that $f(n) \leqq 2^{2^n}$, but no sharper bound seems to have been known. In this paper we obtain easily a sharper bound for $f(n)$, namely, $2^{n(n-1)}$. We then pursue the argument and prove a theorem saying that $f(n)$ is precisely the number of certain $n \times n$ matrices of zeros and ones. The counting of these matrices seems to be well adapted to the methods of work on a computer. With the help of a computer it was then found by Mr. Elio Marzullo that $f(4) = 355$. In the process of the counting and the listing of these matrices and therefore of these topologies, we hit upon the concept of an *n-basic number*. It is not clear,

at least to the author, what interesting properties these n-basic numbers have or can be expected to have.

2. THEOREM. *On the finite set E of n elements there are as many topologies (not taking homeomorphisms into account) as there are $n\times n$ matrices (a_{ij}) of zeros and ones with $a_{ii}=1$ for all i and with the following property*:

(*): *For every ordered pair (R_i, R_j) $(i, j=1, 2, \cdots, n)$ of rows of (a_{ij}) and for every index k, $a_{ji}=1=a_{ik}$ implies $a_{jk}=1$.*

3. *Proof of the Theorem.* First we note that the number of filters on the set E is 2^n-1; for, each such filter is generated by one nonempty subset of E. We can therefore specify one such filter by specifying the basis, namely the set which generates it, thus: $\mathcal{H}_S=$ the filter generated by $S\subset E$, which is the same as the class of all supersets of S in E. If τ is a topology on E, then a family of τ-neighborhoods for $x\in E$ has to be a filter on E, containing x. So every family of neighborhoods of $x\in E$ is specified by a filter $\mathcal{H}_{S(x)}$ where $S(x)$ is a certain subset of E containing x. Given $x\in E$ there are 2^{n-1} possible choices for $S(x)$. So there are 2^{n-1} possible choices for a filter of neighborhoods of x. Hence the number of possible choices for the neighborhood system of the whole space E cannot be greater than $2^{n(n-1)}$. This number therefore is an upper bound for $f(n)$.

Not all these $2^{n(n-1)}$ choices actually give rise tod ifferent topologies on E. Let $\mathcal{H}_{S(x_i)}$ be the filter generated by the set $S(x_i)$. This filter as well as its generator $S(x_i)$ are uniquely specified by the row vector $(a_{i1}, a_{i2}, \cdots, a_{in})$ where $a_{ij}=1$ if $x_j\in S(x_i)$ and $a_{ij}=0$ if $x_j\notin S(x_i)$. So any choice of a possible basis of neighborhoods for a topology τ on E can now be described by n row vectors $S(x_i)$, $i\in N$ and hence by the matrix (a_{ij}) of zeros and ones, with all the elements in the leading diagonal being one. The ith row of the matrix represents a subset of E which generates a filter of neighborhoods of x_i. Hence, $S(x_i)$, being the intersection of all τ-neighborhoods of x_i, has to be τ-open. Thus the system $\mathbf{S}=\{S(x_i), i\in N\}$ and therefore the matrix (a_{ij}) associated with $\mathbf{S}$ as above, give us τ-open sets. We claim that in order that such an $\mathbf{S}$, together with the empty set be actually a basis of τ-open sets it is necessary and sufficient that the matrix (a_{ij}) has property (*).

To see the necessity part of the assertion, reason as follows: $a_{ji}=1$ simply states that $T_{ij}=S(x_i)\cap S(x_j)$ contains x_i. If, in addition, $a_{ik}=1$ but $a_{jk}=0$ then the latter would mean that T_{ij} contains x_i but not x_k and thus there would be an open set containing x_i but not x_k. This cannot be; for, every open set containing x_i has to contain $S(x_i)$, the generator of the filter of neighborhoods of x_i and $S(x_i)$ contains x_k, in virtue of a_{ik} being 1. Thus (*) is satisfied.

Conversely, to see that (*) suffices to make $\mathbf{S}'=$ the class $\{S\in\mathbf{S}$, and $\varnothing$, the empty set$\}$ define a basis for τ, it is enough to prove that $\mathbf{S}'$ has the property that if A and B belong to $\mathbf{S}'$, then each point of $A\cap B$ is contained in a set C belonging to $\mathbf{S}'$ such that $C\subset A\cap B$. Our hypothesis (*) means that T_{ij} is $S(x_i)$ whenever $S(x_j)$ contains x_i. If $S(x_j)$ does not contain x_i, then there are two cases: either $S(x_i)$ contains x_j or not. In the former case, applying (*) to the ordered

pair (R_j, R_i) we see that $T_{ij}=S(x_j)$. In the latter case T_{ij} is either empty or contains some other x_k. Summing up these arguments, we see that T_{ij} falls into one of these four categories. It is either $S(x_i)$ or $S(x_j)$ or empty or finally it contains some x_k, $k\neq i, j$. We have only to follow up this last possibility.

Let $T=T_{ij}\subset E\setminus\{x_i, x_j\}$. We shall show that, for every k for which $x_k\in T$, the $S(x_k)$ associated with the kth row of (a_{ij}) is contained in T. For this, we exhibit the relevant elements from the ith, jth, and kth rows as follows:

ith row:	1	a_{ij}	a_{ik}	a_{ip}
jth row:	a_{ji}	1	a_{jk}	a_{jp}
kth row:	a_{ki}	a_{kj}	1	a_{kp}

Observe that, in the above, $a_{ij}=0=a_{ji}$, since x_i and x_j are not in T. Also $a_{ik}=1=a_{jk}$, since $x_k\in T$. In order to prove our assertion we have only to show that whenever there is an index $r(\neq k)\in N$ such that $x_r\notin T$, then $a_{kr}=0$. Now $x_r\notin T$ means that x_r does not belong to at least one of $S(x_i)$ or $S(x_j)$. Let us take $x_r\notin S(x_i)$. The other possibility can be similarly handled. Write $r=p$. Thus $a_{ip}=0$. Now if $a_{kp}=1$, then, applying (*) to the ordered pair (R_k, R_i), we see that $a_{ik}=1=a_{kp}$ gives $a_{ip}=1$ which is not true. Hence a_{kp} has to be zero. Thus we have proved that, for every $r\neq k$, $x_r\notin T$ implies $a_{kr}=0$. In particular, of course, $a_{ki}=0=a_{kj}$. This completes the proof that for every $x_k\in T=T_{ij}$, T contains the set $S(x_k)$ belonging to $\mathbf{S}'$.

The above completes the justification of the claim that $\mathbf{S}'$ is a basis of τ-open sets for E if and only if the matrix (a_{ij}) associated with $\mathbf{S}$ has property (*). Thus with each topology on E we can associate a unique matrix (a_{ij}) with property (*) and with each such matrix we can associate a "valid" topology on E. From the fact that this association makes $\mathbf{S}'$ a basis of open sets it is also clear that two different matrices with property (*) lead to two distinct (not necessarily nonhomeomorphic) topologies on E. The theorem therefore stands proved.

4. In counting the matrices with property (*) we proceed as follows: First delete the main diagonal, which is just an n-string of ones. We then have n rows of $(n-1)$-vectors, each entry in these vectors being zero or one. Write these n rows of zeros and ones in succession as if it were a single number written in binary scale. In this form it represents an integer less than $2^{n(n-1)}$. Each of these numbers corresponds in a 1-1 manner to a "valid" topology on E. So we make the following

DEFINITION. *An n-basic number is a positive integer not greater than $2^{n(n-1)}$, which leads to a matrix of zeros and ones, having ones in all the entries of the leading diagonal, and which possesses property* (*), *provided the following operations are performed on the number*:

(a) *Express the number in binary scale.*

(b) *Add enough zeros to the left to bring the total number of digits to $n(n-1)$.*

(c) *Break this binary number from left to right into n blocks of $(n-1)$ entries each.*

(d) *Write the blocks one after another in the form of n rows of* $(n-1)$*-vectors to form an* $n\times(n-1)$ *matrix.*

(e) *Remembering that these should be the off-diagonal elements of an* $n\times n$ *matrix, supply the leading diagonal of ones by writing ones in the "proper" places.*

CONVENTION. The n-basic number corresponding to the discrete topology on E would be in the spirit of our definition, the number corresponding to the matrix (δ_{ij}), (δ_{ij} denoting Kronecker delta) and so this number would have to be zero. But we shall make a convention of taking this number to be $2^{n(n-1)}$. This convention has been followed in the counting process but this may not be necessary for a study of n-basic numbers.

REMARK. Note that the trivial topology on E always corresponds to the number $2^{n(n-1)}-1$.

5. We give below the results of the listing of these n-basic numbers for $n=2$, 3. Each of these numbers defines a distinct topology on the set E of n elements and from our theorem it follows that there are no more. The process of programming and of counting these topologies was performed on an IBM 7094 computer with a straightforward Fortran Programming by Mr. Elio Marzullo of the Department of Physics, University of Illinois, Urbana, to whom my thanks are due. A list of the 355 4-basic numbers may be had, on request, from him or from the author.

$n=2$	1	2	3	4					
$f(2)=4$									
$n=3$	1	2	3	4	5	8	10	11	12
$f(3)=29$	14	15	16	18	20	28	30	32	33
	35	40	43	48	49	51	52	53	60
	63	64							

6. Among the questions that may be asked about n-basic numbers, we draw the attention of the reader to the following, which may probably lead to more knowledge about an expression for $f(n)$: What are necessary and sufficient conditions in order that an n-basic number be also an $(n+1)$-basic number?

This work was supported by NSF Grant GP1909.

Author's present address: Birla Institute of Technology and Science, Pilani, (Rajasthan) India.

Said a young man named A. Grothendieck:
"In Geometry I'm rather weak.
I'm no Altshiller-Court;
It's just not my forte,
So I'd best make it more *Algébrique*,"

ARISTOTLE NICHOLAS ONYMOUS, Harvard University

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

ON A METHOD OF CONSTRUCTION OF PARTIAL GEOMETRIES AND PARTIAL BOLYAI-LOBACHEVSKY PLANES

ESTHER SEIDEN, Michigan State University

Dedicated to JERZY NEYMAN on the occasion of his 70th birthday.

1. Summary. The concept of partial geometries was introduced by R. C. Bose [3]. The method of construction of partial geometries described here was not mentioned by Bose. It applies only to projective planes of even order, say $2h$, provided that they do possess an oval of size $2h+2$. It is pointed out here that such planes can be split into two partial geometries. One of them consists of all secants, the lines which include two points of the oval, after removing the points of the oval. The other partial geometry is formed by the remaining lines of the projective plane i.e. by the lines which do not contain any point of the oval, nonintersectors. The duals of these partial geometries are again partial geometries. L. M. Graves [4] introduced the concept of finite Bolyai-Lobachevsky planes and constructed some examples of such planes. Graves states, "It would be interesting to know methods for construction of additional examples of finite B-L planes (in particular homogeneous planes) and learn the limitations on the number n of points on a line and number ν of lines on a point." This request was answered to some extent by R. Sandler [7] and T. G. Ostrom [6]. Both gave methods of construction of infinite classes of finite B-L planes, but their planes are not homogeneous in the sense of Topel [9], according to which all the lines have to contain the same number of points. Topel proved that there is no homogeneous B-L plane with finite lines containing more than two points. It is especially interesting, therefore, to construct finite partial B-L planes and find out the limitations on the number of points on a line. The class of B-L partial planes constructed here are homogeneous in the sense of Topel. They are also homogeneous in the sense of Graves, i.e. "that for each pair of points there is a collineation carrying the first into the second."

2. Construction of partial geometries. Bose, in [3] defines a partial geometry (r, k, t) as a system of undefined points and lines and an undefined relation "incidence" satisfying axioms A1 through A4 below:

A1. *Any two points are incident with no more than one line.*

A2. *Each point is incident with r lines.*

A3. *Each line is incident with k points.*

A4. *If a point P is not incident with the line l, there pass through P exactly t lines ($t \geqq 1$) intersecting l.*

Suppose now that there exists a projective plane of order $2h$ which possesses an oval consisting of $2h+2$ points. One can classify the lines of such a plane into two categories. The first category consists of lines, henceforth called secants, including two points of the oval. The second category consists of lines, henceforth called nonintersectors, not including any points of the oval. If we remove now the points of the oval from the lines of the first category then each of the two categories of lines separately forms a partial geometry. Each of these partial geometries include all the points of the original projective plane except of the points of the oval. Hence there are $4h^2-1$ points in each of the partial geometries. Two points of the oval determine a unique secant. The number of lines of the partial geometry formed by the secants will therefore be $(h+1)(2h+1)$. The remainder of the lines of the projective plane, $2h^2-h$ in number, form the second partial geometry. Each point in the projective plane which does not belong to the oval is an intersection of $h+1$ secants and h nonintersectors. Hence r for the two partial geometries will be equal to $h+1$ and h respectively. The number of points on a line k is equal to $2h-1$ and $2h+1$ respectively, because from each secant two points belonging to the oval are removed. Finally t, the number of lines passing through any point X not on a line l which intersect l, is equal to $h-1$ and h respectively, since through X there pass $h+1$ secants. Two of them intersect l in the removed points of the oval. The remaining $h-1$ secants and the $h+1$ nonintersectors intersect l.

There are infinitely many realizations of partial geometries of this kind. It is enough to observe that an Arguesian plane of order $2h=2^m$, m a positive integer, possesses an oval consisting of $2h+2$ points. As to the non-Arguesian planes of even order it is not known whether or not they possess an oval consisting of $2h+2$ points. Hence no partial geometries based on them could be exhibited by the described method.

REMARKS. These two partial geometries were discussed by the author in terms of partially balanced block designs in [8]. The duals of the partial geometries formed by the secants are also partially balanced block designs. To describe the parameters of these designs one would have to introduce the language of the designs. It will not be done here. The duals of the partial geometries formed by the nonintersectors form resolvable balanced block designs. This was observed by the author independently and reported to R. C. Bose. Bose replied that he made this observation long ago and used it to prove a theorem in [2]. It may be pointed out in this connection that in the proof of this paper the oval consisted of a conic and its center but it applies equally well to any oval.

3. Construction of Bolyai-Lobachevsky planes. Graves defines a plane $\mathcal{P}$ to be a finite B-L plane if it satisfies the following axioms:

A1. *The plane $\mathcal{P}$ is a finite collection of elements called points.*
A2. *There are certain distinguished subsets of the plane $\mathcal{P}$ called lines.*
A3. *There are at least two points on each line.*
A4. *Two distinct points lie on one and only one line.*

A5. *The plane* $\mathcal{P}$ *contains at least four points, no three of which lie on one line.*

A6. *If a subset of* $\mathcal{P}$ *contains three points not on a line, and contains all the lines through any pair of its points, then the subset contains all the points of* $\mathcal{P}$.

A7. *Through each point* X *not on a line* b *there pass at least two lines not meeting* b.

It will be shown here that the dual of the partial geometry formed from the nonintersectors of the projective plane of even order is a partial B-L plane. The number of points of this dual plane is clearly $2h^2-h$, while the number of lines is $4h^2-1$. The number of points on a line is $n=h$, and the number of lines through a point is $\nu=2h+1$. Axioms A1 through A4 are trivially satisfied. Assume now that $h\geqq 2$; then axiom A5 will be satisfied, since it is easy to see that the partial geometry of the nonintersectors contains at least four lines, no three concurrent. Take any point in this partial geometry and two lines through it. Consider two distinct points on one of these lines different from the first chosen point. Such points do exist because $k=2h+1$. Clearly, there must exist two lines passing respectively through each of these points which do not intersect on the second line through the first chosen point. These four lines will form a set of lines, no three concurrent. It will be shown next that if $h>3$ then A6 will be satisfied. Let a subset of $\mathcal{P}$ contain three points not on one line, say, A, B, C. Consider the lines through A. There are $2h+1$ such lines. Two of these lines are the lines AB and AC. The premise implies that the line BC belongs to the subset. This line must include $h-2$ additional points which belong one by one to some additional $h-2$ lines through A. Thus the subset already includes $h(h-1)$ points different from A belonging to h lines through A. Notice further that the $2h+1$ lines through A include all the points of $\mathcal{P}$ because of axiom A4. Now suppose that there exists a point of $\mathcal{P}$ which does not belong to the subset. This point must belong to one of the lines through A other than the h lines counted previously. Through this point must pass $2h$ additional lines and none of them can contain more than one of the $h(h-1)$ points which already belong to the subset, thus $2h\geqq h(h-1)$. This is clearly impossible unless $h\leqq 3$. Axiom A7 will be satisfied for the following reason. If a point X is not on a line b then there will be h distinct lines joining X to each of the points of b. Each of these lines has $h-2$ more points. Thus the total number of points through X incident with lines intersecting b is $h(h-1)$. Since there are $2h^2-h$ points in the plane, X must be joined to $2h^2-h-1$ points. Thus X will be joined to the remainder of h^2-1 points by lines parallel to b. This gives $h+1$ lines passing through X and parallel to b. Notice that the number of parallel lines is independent of the choice of X and b.

4. The homogeneity of partial B-L planes. It will be shown presently that the partial B-L planes obtained from Arguesian planes are homogeneous in the sense of Graves [4], provided that the oval used for their construction consists

of a conic and its center, i.e. the point of intersection of the tangents to the conic. The proof follows from some simple properties of the Arguesian planes of even order, say, 2^m.

R. C. Bose showed in [1] that, if the 2^m+1 points of a nondegenerate conic are represented by the equations $ax^2+by^2+cz^2+fyz+gxz+hxy=0$, then the (2^m+2)-nd point of the oval, the center of the conic can be represented by the coordinates (f, g, h). We may conclude further that three points of a conic together with its center determine the conic uniquely.

Another property of Arguesian planes to be used here is that their collineations are transitive on quadrilaterals (see [5]).

A collineation which leaves the oval fixed will clearly be a collineation of the partial planes formed by the nonintersectors and hence a collineation of their duals, the partial B-L planes. One can obtain such a collineation by mapping a quadrilateral consisting of three points of a conic and its center into a quadrilateral of the same type in respect to the same conic. Since three points of a conic can be chosen in $(2^{2m}-1)2^m$ ways, the collineation group will be of that order. The question which remains to be answered is whether this group contains a collineation which maps any arbitrary point not belonging to the oval A into any other arbitrary point of the same kind B. To see that the answer is in affirmative it is enough to join the points A and B to the center of the conic, say C. The lines AC and BC, not necessarily distinct, include another point of the conic. Consider two more distinct secants through A and B respectively. The collineation which maps the quadrilateral consisting of C, the point of the conic on AC and the two points of the conic on the secant through A into the corresponding quadrilateral through B will map A into B. This concludes the proof.

Acknowledgment. It is a pleasure to acknowledge the stimulating discussions with H. Rubin during the preparation of this paper.

Supported in part by Office of Naval Research Grant No. NONR-2587(02).

References

1. R. C. Bose, Mathematical theory of factorial design, Sankhyā, 8 (1947) 107–165.
2. R. C. Bose and S. S. Shrikhande, On the construction of sets of mutually orthogonal latin squares and the falsity of a conjecture of Euler, Trans. Amer. Math. Soc., 95 (1960) 191–209.
3. R. C. Bose, Strongly regular graphs, partial geometries and partially balanced designs, Pacific J. Math. 13 (1963) 389–419.
4. L. M. Graves, A Finite Bolyai-Lobachevsky plane, this MONTHLY, 69 (1962) 130–132.
5. Marshall Hall, Jr., The Theory of Groups, Macmillan, New York, 1959.
6. T. G. Ostrom, Oval and finite Bolyai-Lobachevsky planes, this MONTHLY, 69 (1962) 899–901.
7. R. Sandler, A finite homogenous Bolyai-Lobachevsky planes, this MONTHLY, 70 (1963) 853–855.
8. E. Seiden, On a geometrical method of construction of partially balanced designs with two associate classes. Ann. Math. Statis., 32 (1961) 1177–1180.
9. B. J. Topel, Bolyai-Lobachevsky planes with finite lines, Rep. Math Colloquium, Ser. 2, vol 5–6, Notre Dame, Ind. (1944) 40–42.

SOME LIMITS INVOLVING BINOMIAL COEFFICIENTS

D. R. HAYES, University of Tennessee

1. In a recent paper [1], H. W. Gould showed that

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=0}^{n}\log\binom{pn}{pk}=\frac{p}{2},\tag{1}$$

where p is a positive integer and $\binom{pn}{pk}$ is a binomial coefficient. The aim of this paper is to determine limits involving sums of the form

$$S_{pq}(x,n)=\sum_{k=0}^{n}x^k\log\binom{pn}{qk}\tag{2}$$

as $n\to\infty$, where x is a given complex number and p and q are positive integers such that $q\leqq p$. The author is indebted to Professor Gould for suggesting this problem.

It turns out that the sequence $\{S_{pq}(x,n)\}_{n>0}$ behaves rather differently according to whether the point x lies in the interior, in the exterior, or on the boundary of the unit disk. When x does not lie on the unit circle, one can answer satisfactorily the questions: (1) What is the "natural" elementary function $\lambda(n)$ having the property that $\lambda(n)\cdot S_{pq}(x,n)$ converges to an interesting limit as $n\to\infty$? and (2) What is that limit? The results are:

(a) If $|x|<1$, then

$$\lim_{n\to\infty}\frac{S_{pq}(x,n)}{\log n}=\frac{qx}{(1-x)^2}.\tag{3}$$

(b) If $|x|>1$ and $p=q$, then

$$\lim_{n\to\infty}\frac{S_{qq}(x,n)}{x^n\log n}=\frac{qx}{(x-1)^2}.\tag{4}$$

(c) If $|x|>1$ and $q<p$, then

$$\lim_{n\to\infty}\frac{S_{pq}(x,n)}{nx^n}=[p\log p-q\log q-(p-q)\log(p-q)]\cdot\frac{x}{x-1}.\tag{5}$$

When $|x|=1$, the sequence $\{S_{pq}(x,n)\}_{n>0}$ can behave rather erratically. For example, the sequence $\{S_{11}(-1,n)/\log n\}_{n>0}$ has exactly two limit points, 0 and $-\frac{1}{2}$. Nevertheless, some results can be obtained. For example, Gould's result (1) deals with $S_{pp}(1,n)$. In section 3 below, we show that if $q<p$, then

$$\lim_{n\to\infty}\frac{S_{pq}(1,n)}{n^2}=\frac{p}{2}+\frac{p^2}{2q}[(1-(q/p))^2\log(1-(q/p))-(q/p)^2\log(q/p)].\tag{6}$$

Professor Gould has also derived this result using the methods of his paper [1]. Furthermore, if ζ is an r-th root of unity different from 1, one can show that

$$\lim_{n\to\infty} \frac{S_{11}(\zeta, rn)}{\log n} = \frac{2\zeta}{(1-\zeta)^2}. \tag{7}$$

The proof of (7) is left to the reader.

2. The following Lemma is useful in proving (3) and (5).

LEMMA. *Let* $\{a(n, k)\}_{n\geqq k}$ *be a "triangular" double sequence of complex numbers such that* $a(n, k)\to a(k)$ *as* $n\to\infty$ *for each* k. *Suppose that* $|a(n, k)| \leqq \theta(k)$ *for all* $n\geqq k$, *where* $\theta(k)$ *is a positive-valued function such that* $\sum_{k=0}^{\infty} \theta(k)x^k$ *converges for* $|x|<1$. *Then for* $|x|<1$, *the series on the right in* (8) *below converges and*

$$\lim_{n\to\infty} \left(\sum_{k=0}^{n} a(n, k)x^k \right) = \sum_{k=0}^{\infty} a(k)x^k. \tag{8}$$

Proof. Since $|a(n, k)| \leqq \theta(k)$ for $n\geqq k$, it is clear that $|a(k)| \leqq \theta(k)$. The convergence of the series on the right in (8) for $|x|<1$ follows from this fact by comparison with the convergent series $\sum_{k=0}^{\infty} \theta(k)x^k$. Now suppose a complex number x with $|x|<1$ and $\epsilon>0$ are given. Choose N so large that $\sum_{k=N+1}^{\infty} \theta(k)|x|^k \leqq \epsilon/4$. Then choose $n>N$ so large that $|a(n, k)-a(k)| \leqq (1-|x|)\cdot(\epsilon/2)$ for all $k\leqq N$. Then we have

$$\begin{aligned} &\left| \sum_{k=0}^{n} a(n, k)x^k - \sum_{k=0}^{n} a(k)x^k \right| \\ &\leqq \sum_{k=0}^{N} |a(n, k) - a(k)| \cdot |x|^k + \sum_{k=N+1}^{n} (|a(n, k)| + |a(k)|)|x|^k \\ &\leqq (1 - |x|)\cdot(\epsilon/2)\cdot \sum_{k=0}^{N} |x|^k + \sum_{k=N+1}^{n} 2\theta(k)\cdot |x|^k \\ &\leqq (1 - |x|)\cdot(\epsilon/2)\cdot \sum_{k=0}^{\infty} |x|^k + 2 \sum_{k=N+1}^{\infty} \theta(k)\cdot |x|^k \\ &\leqq \epsilon/2 + 2\cdot(\epsilon/4) = \epsilon. \end{aligned}$$

Thus, $\lim_{n\to\infty} (\sum_{k=0}^{n} a(n, k)x^k - \sum_{k=0}^{n} a(k)x^k) = 0$, from which the assertion of the Lemma follows easily. This completes the proof.

Proof of (3). For $n\geqq k$, put

$$a(n, k) = \log \binom{pn}{qk} \Big/ \log n.$$

Then

$$a(n, k) = \sum_{i=1}^{qk} \log (pn - i + 1)/\log n - \log (qk)!/\log n. \tag{9}$$

Since the binomial coefficients are integers, it follows from (9) that

$$\begin{aligned}|a(n, k)| &\leqq \sum_{i=1}^{qk} \log (pn - i + 1)/\log n \\ &\leqq (qk) \cdot \frac{\log (pn)}{\log n} = (qk)\left(1 + \frac{\log p}{\log n}\right) \\ &\leqq (qk)(1 + \log p/\log 2) \quad \text{for } n > 1.\end{aligned}$$

Also, from (9)

$$\lim_{n\to\infty} a(n, k) = \sum_{i=1}^{qk}\left(\lim_{n\to\infty} (\log (pn - i + 1)/\log n)\right) = \sum_{i=1}^{qk} 1 = qk.$$

Thus, if we take $\theta(k) = (1 + \log p/\log 2)qk$ and $a(k) = qk$, then the Lemma applies to the sequence $a(n, k)$ and yields, for $|x| < 1$,

$$\lim_{n\to\infty} \frac{S_{pq}(x, n)}{\log n} = \sum_{k=0}^{\infty} (qk)x^k = q\sum_{k=0}^{\infty} kx^k = \frac{qx}{(1 - x)^2},$$

which is (3).

Proof of (4). We have

$$\begin{aligned}S_{qq}(x, n) &= \sum_{k=0}^{n} x^k \log \binom{qn}{qk} = x^n \sum_{k=0}^{n} x^{k-n} \log \binom{qn}{qk} \\ &= x^n \sum_{s=0}^{n} x^{-s} \log \binom{qn}{qn - qs} = x^n \sum_{s=0}^{n} \left(\frac{1}{x}\right)^s \log \binom{qn}{qs} \qquad (10) \\ &= x^n S_{qq}\left(\frac{1}{x}, n\right),\end{aligned}$$

where $s = n - k$. Thus, if $|x| > 1$ we get

$$\lim_{n\to\infty} \frac{S_{qq}(x, n)}{x^n \log n} = \lim_{x\to\infty} \frac{S_{qq}(1/x, n)}{\log n} = \frac{q(1/x)}{(1 - (1/x))^2} = \frac{qx}{(x - 1)^2}$$

by (3). This completes the proof.

Proof of (5). Suppose $|x| > 1$. As in (10), we find that

$$\frac{S_{pq}(x, n)}{nx^n} = \sum_{k=0}^{n} \left(\frac{1}{x}\right)^k \log \binom{pn}{q(n - k)} \Big/ n.$$

Our aim is to apply the lemma with $a(n, k) = \log \binom{pn}{q(n-k)}/n$. We have, clearly $|a(n, k)| \leqq (\log (1+1)^{pn})/n = (pn) \log 2/n = p \log 2$, so that $\theta(k) = p \log 2$ meets the requirements of the lemma. To evaluate the $a(k)$, we need the following familiar limit:

$$\lim_{r\to\infty} \left(\frac{r!^{1/r}}{r}\right) = e^{-1}. \qquad (11)$$

Since $a(n, k)=\log \binom{pn}{q(n-k)}^{1/n}$, in order to evaluate $a(k)$, it suffices to find $\lim_{n\to\infty} \binom{pn}{q(n-k)}^{1/n}$. This may be accomplished as follows. For brevity put $[r]=(r!^{1/r})/r$, so that by (11) $\lim_{r\to\infty} [r]=e^{-1}$. Then

$$\begin{aligned}\binom{pn}{q(n-k)}^{1/n} &= \left(\frac{(pn)!}{((p-q)n+qk)!(q(n-k))!}\right)^{1/n}\\ &= [pn]^p([(p-q)n+qk])^{q-p-(qk/n)}\\ &\quad\times([q(n-k)])^{(qk/n)-q}p^p\cdot(p-q+(qk/n))^{q-p-(qk/n)}\\ &\quad\times(q-(qk/n))^{(qk/n)-q},\end{aligned}$$

so that

$$\lim_{n\to\infty}\binom{pn}{q(n-k)}^{1/n} = (e^{-p}\cdot e^{p-q}\cdot e^{q})p^p(p-q)^{q-p}q^{-q} = p^pq^{-q}(p-q)^{q-p}.$$

Therefore, $a(k)=p\log p-q\log q-(p-q)\log(p-q)$ for every value of k. We are now in a position to apply the lemma. We find that

$$\begin{aligned}\lim_{n\to\infty}\frac{S_{pq}(x,n)}{nx^n} &= \sum_{k=0}^{\infty}[p\log p-q\log q-(p-q)\log(p-q)]\left(\frac{1}{x}\right)^k\\ &= [p\log p-q\log q-(p-q)\log(p-q)]\cdot\frac{x}{x-1}\cdot\end{aligned}$$

This completes the proof of (5).

3. The equation (6) can be established in the following way:

$$\begin{aligned}\frac{S_{pq}(1,n)}{n^2} &= \frac{1}{n^2}\sum_{k=0}^{n}\log\binom{pn}{qk}\\ &= \frac{1}{n^2}\sum_{k=1}^{n}\log\left[\left(\frac{pn}{pn-qk}\right)\cdot\prod_{j=1}^{qk}\frac{pn-j}{j}\right]\\ &= \frac{1}{n^2}\sum_{k=1}^{n}\log\left(\frac{pn}{pn-qk}\right)+\frac{1}{n^2}\sum_{k=1}^{n}\sum_{j=1}^{qk}\log\left(\frac{p-(j/n)}{(j/n)}\right)\\ &= S_1+S_2.\end{aligned}\tag{12}$$

One can show easily enough that $S_1\to 0$ as $n\to\infty$. And S_2 is recognized as an approximating sum to the double integral of the function $\log((p-y)/y)$ over the triangular region $A=\{(x,y)\mid 0\leq x\leq 1 \text{ and } 0\leq y\leq qx\}$. Therefore we have

$$\lim_{n\to\infty}\frac{S_{pq}(1,n)}{n^2} = \iint_A \log\left(\frac{p-y}{y}\right)dA.$$

The double integral may be evaluated by standard methods and yields (6).

Reference

1. H. W. Gould, Sums of logarithms of binomial coefficients, this MONTHLY, 71 (1964) 55–58.

ON ZEROS OF A CONTINUOUS NOWHERE DIFFERENTIABLE FUNCTION

J. S. Lipinski, University of Łódz, Poland

In the note [2], of S. Roy Schubert, the author asks whether there exists a continuous mapping $f: I\to R$ and a set $A\subset I$ such that f is nowhere differentiable, $f=0$ on A, and $m(A)>0$. The letter I denotes here the closed interval $\langle 0, 1\rangle$, and m the Lebesgue measure on the real line R. The answer given by Schubert is affirmative: "Given $\epsilon>0$, it is possible to construct such a function f on I with the corresponding $A\subset I$ that $m(A)>1-\epsilon$." Unfortunately, the function f introduced in [2] does not have the announced properties. His f is continuous and the set of its zeros is of the desired measure, but as we shall show, the function f is differentiable almost everywhere in A.

The function f is constructed as follows: Let W be a continuous nowhere differentiable function defined on I such that $W(0)=W(1)=0$. In particular Schubert chooses W to be a well-known function of van der Waerden (see [3]); hence $W(t)>0$ for $0<t<1$. Let $\epsilon>0$ be given and let C_ϵ denote a Cantor-like set on I such that $m(C_\epsilon)>1-\epsilon$. Then $C_\epsilon=I\setminus\bigcup_{n=1}^{\infty} E_n$ where $E_n=(a_n, b_n)$. Set $f=0$ on C_ϵ and define f to be a copy of W on the closure $\overline{E}_n$ of E_n for each n. According to this definition

$$f(x) = (b_n - a_n)W\left(\frac{x-a_n}{b_n-a_n}\right) \quad \text{for } a_n \leqq x \leqq b_n.$$

Set $M=\max_{x\in I} W(x)$, $M_n=\max_{x\in E_n} f(x)$. Then $0<M<1$ and $M_n=(b_n-a_n)M$. Let $B_k=\{x: D^+(x)>1/k,\ x\in C_\epsilon\}$, where D^+ denotes the upper left-hand Dini derivative. Let further $x_0\in B_k$. Then there exists a sequence of positive numbers $h_m\to 0$ such that $h_m^{-1}f(x_0+h_m)>k^{-1}$. Clearly $x_0+h_m\overline{\in} C_\epsilon$. Let $x_0+h_m\in E_{n_m}$. Then

$$\frac{b_{n_m}-a_{n_m}}{b_{n_m}-x_0} > \frac{M_{n_m}}{a_{n_m}-x_0}\cdot\frac{a_{n_m}-x_0}{b_{n_m}-x_0} \geqq \frac{f(x_0+h_m)}{h_m}\cdot\frac{a_{n_m}-x_0}{b_{n_m}-x_0} > \frac{1}{k}\left(1-\frac{b_{n_m}-a_{n_m}}{b_{n_m}-x_0}\right).$$

Hence,

$$\frac{b_{n_m}-a_{n_m}}{b_{n_m}-x_0} > \frac{1}{k+1}\,.$$

Further, in view of

$$\frac{m\{(I\setminus C_\epsilon)\cap(x_0, b_{n_m})\}}{b_{n_m}-x_0} \geqq \frac{b_{n_m}-a_{n_m}}{b_{n_m}-x_0},$$

we have

$$l(x_0) = \limsup_{h\to 0}\frac{m\{(I\setminus C_\epsilon)\cap(x_0, x_0+h)\}}{h} \geqq \frac{1}{k+1}\,.$$

According to the well-known Lebesgue theorem on density points (see e.g. [1] p. 129) for any set C_ϵ, $m\{x: l(x)\neq 0,\ x\in C_\epsilon\}=0$. Thence $m(B_k)=0$. Let $B=\{x: D^+(x)>0,\ x\in C_\epsilon\}$. In view of $B=\bigcup_{k=1}^{\infty} B_k$ we have $m(B)=0$. An analogous argument may be applied to the other Dini derivatives. Thus the set of points $x\in C_\epsilon$ in which at least one Dini derivative is different from zero is a set of the measure zero. In consequence $f'(x)=0$ almost everywhere for $x\in C_\epsilon$.

I shall now show that the question asked in [2] has an affirmative answer. Moreover, I shall not only prove the existence of a continuous nowhere differentiable function with a set of zeros of positive measure, but also I shall present a complete characterization of this set.

From the continuity of the function it results that this set must be closed. The function cannot be constant on any interval, and so the set of zeros does not contain any interval. A set which does not contain any interval is nowhere dense. These two properties characterize completely the set of zeros.

For any closed nowhere dense set $C\subset I$ there exists a continuous nowhere differentiable function f which vanishes for $x\in C$ and is positive for $x\overline{\in}(0, 1)\cap C$.

REMARK. The above theorem gives the answer concerning the question in [2] since for any $\epsilon>0$ there exists in I a closed nowhere dense set of a measure larger than $1-\epsilon$.

Proof. Arrange in a sequence $\{A_n\}$ all intervals $(2^{-i}(k-1),\ 2^{-i}k)$, where $k=1, 2, \cdots, 2^i$; $i=1, 2, \cdots$. Let $(0, 1)\backslash C=\bigcup_{i=1}^{\infty}(a_i, b_i)$ where $(a_i, b_i)\cap(a_j, b_j)=0$ for $i\neq j$. Let F_1 be any interval (a_i, b_i) contained in A_1 in case such intervals exist, and the empty set in case there are no such intervals. Let F_n be any interval (a_i, b_i) contained in A_n, $F_n\neq F_j$ $(j=1, 2, \cdots, n-1)$ in case such an interval exists, and the empty set in case there are no such intervals. Set $f(x)=0$ for $x\in C$,

$$f(x) = m(A_i)W\left(\frac{x - a_{n_i}}{b_{n_i} - a_{n_i}}\right) \quad \text{for } x \in F_i = (a_{n_i}, b_{n_i})$$

and let

$$f(x) = (b_n - a_n)W\left(\frac{x - a_n}{b_n - a_n}\right) \quad \text{for } x \in (a_n, b_n) \neq F_k \quad (k = 1, 2, \cdots).$$

In view of $W(x)>0$ for $0<x<1$ we have $f(x)>0$ for $x\in(a_{n_i}, b_{n_i})$ and $x\in(a_n, b_n)$. Then $\{x: 0<x<1;\ f(x)=0\}=C$. The function $f(x)$ is continuous in (a_n, b_n) and continuous from the right at the points a_n. We shall prove the continuity of this function from the right at other points i.e. in the set $D=C-\bigcup_{n=1}^{\infty}\langle a_n, b_n\rangle$. Let $\epsilon>0$ and $x\in D$. Denote by S the union of intervals $\langle a_n, b_n\rangle$ satisfying the conditions $a_n>x$, $b_n-a_n\geqq\epsilon$ and $m(A_i)\geqq\epsilon$ for $(a_n, b_n)=F_i$. The number of intervals which satisfy the first and second or first and third inequalities is finite, and the union is also a closed set. Put $\delta=\inf_{y\in s}|x-y|$; then $\delta>0$. If $y\in C$ then $f(y)=0$ and $|f(x)-f(y)|=0<\epsilon$. If $y\in(a_j, b_j)=F_i$, then we have $F_i\cap S=0$ and $0<f(y)$

$<m(F_i)<\epsilon$. Finally, if $x\in(a_j, b_j)\setminus\bigcup_{i=1}^{\infty} F_i$ then $0<f(x)<b_j-a_j<\epsilon$. In this way we have proved the right-hand continuity in any point of the interval $\langle 0, 1\rangle$. The proof of left-hand continuity may be performed analogously. In consequence f is continuous in I.

There remains to prove the nondifferentiability of f. Nondifferentiability in the points of the interval $\langle a_n, b_n\rangle$ is clear. Let $x\in C\setminus\bigcup_{n=1}^{\infty}\langle a_n, b_n\rangle$ and let $\{A_{n_i}\}$ be any decreasing sequence which contains the points x. Every interval A_{n_i} contains infinitely many intervals (a_n, b_n), and thus the set F_{n_i} is not empty. The function f attains in F_{n_i} a maximum equal to $M\cdot m(A_{n_i})$. Suppose that this maximum is attained in the point x_i. We have $\lim_{i\to\infty} x_i=x$ and $(f(x_i)-f(x))/|x_i-x|\geqq Mm(A_{n_i})/m(A_{n_i})=M$. Denote by y_j one of the end-points of F_{n_j}. Then $\lim_{j\to\infty} y_j=x$ and $(f(y_j)-f(x))/|y_j-x|=0$. Thus we have proved that the function f is not differentiable in the point x, which completes the proof.

References

1. S. Saks, Theory of the Integral, Warszawa 1939.
2. S. Roy Schubert, On a Function of van der Waerden, this MONTHLY, 70 (1963) 402.
3. B. L. van der Waerden, Ein einfaches Beispiel einer nicht differenzierbaren stetigen Funktion, Math. Zeitschrift, 32 (1930).

SOME LIMITS INVOLVING BINOMIAL COEFFICIENTS

L. CARLITZ, Duke University

1. Put $S_{pq}(x, n)=\sum_{k=0}^{n} x^k \log \binom{pn}{qk}$, where p, q are positive integers, $p\geqq q$, and x is a given complex number. D. R. Hayes [2] has investigated the behavior of $S_{pq}(x, n)$ as $n\to\infty$. He showed that the sequence $\{S_{pq}(x, n)\}$ behaves rather differently according as x lies in the interior, in the exterior or on the boundary of the unit circle. In particular, he showed that if $|x|<1$ then

$$\lim_{n=\infty} \frac{S_{pq}(x, n)}{\log n}=\frac{qx}{(1-x)^2}. \tag{1}$$

When $x=1$, $\lim n^{-2}S_{pq}(1, x)$ is evaluated. Earlier H. W. Gould [1] had discussed some special cases.

Hayes pointed out that when $|x|=1$ the situation is rather different. In particular he stated the formula

$$\lim_{n=\infty} \frac{S_{11}(\zeta, rn)}{\log n}=\frac{2\zeta}{(1-\zeta)^2}, \tag{2}$$

where $\zeta^r=1$, $\zeta\neq 1$.

In the present note we show that if $|x|=1$, $x\neq 1$, then

$$S_{pp}(x, n)=\frac{px(1+x^n)}{(1-x)^2}\log n+O(1), \tag{3}$$

$$(4)\qquad S_{pq}(x,n) = -\frac{nx^{n+1}}{1-x}\left[p\log p - q\log q - (p-q)\log(p-q)\right] + O(\log n)$$

$$(p > q).$$

2. The identity (with $a_0=0$)

$$\sum_{k=0}^{n} a_k x^k = -\frac{x}{1-x}\sum_{k=0}^{n-1} x^k(a_k - a_{k+1}) - \frac{x^{n+1}a_n}{1-x}$$

gives

$$(5)\qquad S_{pq}(x,n) = -\frac{x}{1-x}\sum_{k=0}^{n-1} x^k \log\binom{pn}{qk}\binom{pn}{q(k+1)}^{-1} - \frac{x^{n+1}}{1-x}\log\binom{pn}{qn}.$$

We have

$$(6)\qquad \sum_{k=0}^{n-1} x^k \log\binom{pn}{qk}\binom{pn}{q(k+1)}^{-1} = \sum_{k=0}^{n-1} x^k \log\prod_{j=1}^{q}\frac{qk+j}{(p-q)n+qk+j}$$

$$-\sum_{k=0}^{n-1}(x^{n-k-1}-x^k)\log\prod_{j=1}^{q}((p-q)n+qk+j) = S_1 - S_2$$

say. By partial summation

$$\sum_{k=0}^{n-1} x^k \log\prod_{j=1}^{q}\frac{qk+j}{(p-q)n+qk+j}$$

$$= \sum_{k=0}^{n-2}\frac{1-x^{k+1}}{1-x}\log\prod_{j=1}^{q}\frac{qk+j}{(p-q)n+qk+j}\,\frac{(p-q)n+qk+q+j}{qk+q+j}$$

$$+\frac{1-x^n}{1-x}\log\prod_{j=1}^{q}\frac{qn-q+j}{pn-q+j}.$$

Since

$$\log\frac{qk+j}{qk+q+j} = \log\left(1-\frac{q}{qk+q+j}\right) = O\left(\frac{1}{k}\right)\qquad (k\to\infty)$$

$$\log\frac{(p-q)n+qk+q+j}{(p-q)n+qk+j} = \log\left(1+\frac{q}{(p-q)n+qk+j}\right) = O\left(\frac{1}{n}\right)\ (p>q)$$

$$\log\prod_{j=1}^{q}\frac{qn-q+j}{pn-q+j} = O(1),$$

it follows that

$$(7)\qquad S_1 = O(\log n)\qquad (p>q).$$

Turning to S_2, we put

$$T(n, k) = \sum_{j=0}^{k} (x^{n-j-1} - x^j) = - \frac{1 - x^{k+1} - x^{n-k-1} + x^n}{1 - x}.$$

Then by partial summation

$$S_2 = \sum_{k=0}^{n-2} T(n, k) \log \prod_{j=1}^{q} \frac{(p - q)n + qk + j}{(p - q)n + qk + q + j}.$$

Now

$$\sum_{k=0}^{n-2} \log \prod_{j=1}^{q} \frac{(p - q)n + qk + q + j}{(p - q)n + qk + j} = \sum_{j=1}^{q} \log \frac{pn - q + j}{(p - q)n + j}.$$

Clearly this sum is $O(\log n)$ when $p > q$ and is equal to $p \log n + O(1)$ when $p = q$. On the other hand, the sum

$$\sum_{k=0}^{n-2} (x^{k+1} + x^{n-k-1}) \log \prod_{j=1}^{q} \frac{qk + j}{qk + q + j} = O(1)\cdot$$

It follows that

$$S_2 = \begin{cases} O(\log n) & (p > q) \\ p \dfrac{1 + x^n}{1 - x} \log n + O(1) & (p = q). \end{cases} \tag{8}$$

Finally, by Stirling's formula,

$$\log \binom{pn}{qn} = [p \log p - q \log q - (p - q) \log (p - q)]n + O(\log n) \quad (p > q). \tag{9}$$

Substituting from (6), (7), (8) and (9) in (5) we get (3) and (4).

Supported in part by NSF Grant GP-1593.

References

1. H. W. Gould, Sums of logarithms of binomial coefficients, this MONTHLY, 71 (1964) 55–58.
2. D. R. Hayes, Some limits involving binomial coefficients, this MONTHLY, 73 (1966) 162–165.

Editorial Note. We have been informed by J. G. Ceder that the four open problems posed by Solomon Marcus at the end of his note, "Open everywhere discontinuous functions," this MONTHLY, 72 (1965) 993–995, are answered affirmatively by example 6.1, p. 105 of the paper entitled "Darboux continuity," by A. Bruckner and J. Ceder, Jahresbericht der Deutschen Mathematiker-Vereinigung, 67 (1965) 93–117.

UNIFORM DISTRIBUTION OF POLYNOMIALS MODULO m

S. R. Cavior, State University of New York at Buffalo

Let f be a polynomial with integral coefficients. We say that the infinite sequence $S(f)=\{f(i): i=1, 2, \cdots\}$ is *uniformly distributed modulo m* (or u.d. (mod m)) if

$$\lim_{n\to\infty} \frac{1}{n} A(n, j, m) = \frac{1}{m},$$

where $A(n, j, m)$ denotes the number of terms among $f(1), f(2), \cdots, f(n)$ that satisfy the congruence

$$f(i) \equiv j \pmod{m}.$$

In [1] Zane noted that the sequence $S(f)$ is u.d. (mod m) if and only if the integers $f(1), f(2), \cdots, f(m)$ constitute a complete residue system (mod m). From this point of view he determined which monomials are u.d. (mod m). The object of the present paper is to find necessary and sufficient conditions for an arbitrary polynomial to be u.d. (mod m).

Suppose $m=p_1^{e_1}\cdots p_r^{e_r}$. For a polynomial $f(x)$ with integral coefficients to be u.d. (mod m), it is necessary and sufficient that

$$f(x) \equiv i \pmod{m} \qquad (1 \leqq i \leqq m) \tag{1}$$

have a unique solution (mod m). One can easily show that, if $N(i)$ denotes the number of solutions of (1) and $N^{(j)}(i)$ denotes the number of solutions of

$$f(x) \equiv i \pmod{p_j^{e_j}}, \qquad (1 \leqq i \leqq p_j,\ 1 \leqq j \leqq r), \tag{2}$$

then

$$N(i) = \prod_j N^{(j)}(i).$$

Accordingly, (1) has a unique solution (mod m) if and only if (2) has a unique solution (mod $p_j^{e_j}$) for $j=1, 2, \cdots, r$.

Following Hardy and Wright [2, Theorem 123, p. 97], we note that the congruence (2) has a unique solution (mod $p_j^{e_j}$) if and only if $f(x)\equiv i$ (mod p_j) has a unique solution (mod p_j) and

$$f'(a) \not\equiv 0 \pmod{p_j} \qquad (a = 1, 2, \cdots, p;\ e_j > 1).$$

Thus, for a polynomial $f(x)$ with integral coefficients to be u.d. (mod m), it is necessary and sufficient that the congruence $f(x)\equiv i$ (mod p_j) be solvable for $i=1, 2, \cdots, p_j$ and $j=1, 2, \cdots, r$, and that $f'(a)\not\equiv 0$ (mod p_j) for $a=1, 2, \cdots, p$ and $e_j>1$. The first condition is clearly equivalent to requiring that $S(f)$ is u.d. (mod p_j), $j=1, 2, \cdots, r$. Therefore, we have

THEOREM 1. *Let* $m = p_1^{e_1} \cdots p_r^{e_r}$ *and let* $f(x)$ *be a polynomial with integral coefficients.* $S(f)$ *is* u.d. (mod m) *if and only if* $S(f)$ *is* u.d. (mod p_j), $j = 1, 2, \cdots, r$, *and*

$$f'(a) \not\equiv 0 \pmod{p_j} \qquad (a = 1, 2, \cdots, p;\ e_j > 1).$$

It should be noted that, when Theorem 1 is applied to monomials, we obtain the result due to Zane [**1**, Theorem 2, p. 163] which states that the sequence $S(ax^k)$ is u.d. (mod m) if and only if $S(ax^k)$ is u.d. (mod p_j) for each prime divisor p_j of m and m is square free. To show that Zane's result is in fact a special case of Theorem 1, we observe that, since

$$kx^{k-1} \equiv 0 \pmod{p_j}$$

has the solution 0, Theorem 1 requires that $e_j = 1$ for $j = 1, 2, \cdots, r$. In other words, m must be square-free.

In the event that m is square-free, Theorem 1 can be simplified in the following way.

THEOREM 2. *Let* $m = p_1 p_2 \cdots p_r$, *where the primes are distinct, and let* $f(x)$ *be a polynomial with integral coefficients.* $S(f)$ *is* u.d. (mod m) *if and only if* $S(f)$ *is* u.d. (mod p_j), $j = 1, 2, \cdots, r$.

As a simple illustration of Theorem 2, suppose that $f(x) = x^3$. Since $S(f)$ is u.d. (mod 5) and u.d. (mod 11), $S(f)$ is u.d. (mod 55).

It might be of interest to note the following result which is a rather simple consequence of Theorem 1.

THEOREM 3. *If* $S(f)$ *is* u.d. (mod $p_1^{e_1} \cdots p_r^{e_r}$) *and* $e_i > 1$ *for* $i = 1, 2, \cdots, r$, *then* $S(f)$ *is* u.d. (mod $p_1^{f_1} \cdots p_r^{f_r}$), *where* $f_i \geqq e_i$ *for* $i = 1, \cdots, r$.

References

1. B. Zane, Uniform distribution modulo m of monomials, this MONTHLY, 71 (1964) 162–164.
2. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, 4th ed., Oxford Univ. Press, New York, 1960.

A THEOREM ON LINEAR SUMMABILITY

D. F. DAWSON, North Texas State University

If M is a set of complex numbers, let M' denote the set of all convergent sequences whose terms belong to M. In [1] we used only the linearity of triangular matrices, regarded as sequence-to-sequence transformations, in order to show that if C is the Cantor set and A is a complex triangular matrix which is convergence preserving over C', then A is convergence preserving. The following theorem includes this result.

THEOREM 1. *If* M *is a set of real numbers, then the following two statements are equivalent:* (1) *if* A *is a complex matrix which is convergence preserving over* M', *then* A *is convergence preserving, and* (2) M *has a finite limit point.*

The proof of Theorem 1 will be given later. Meanwhile, we will show that Theorem 1 is the best possible in the sense that "complex matrix" in (1) cannot be replaced with "linear sequence-to-sequence transformation." Let R denote the set of rationals.

THEOREM 2. *There exists a linear sequence-to-sequence transformation A which has the following properties*: (1) *the domain of A is the set of all real sequences*, (2) *A is regular over R'*, (3) *A is not convergence preserving.*

Proof. Let X denote a basis of R' and let X^* denote the finite linear completion of X using real coefficients. Let Y denote the complement of X^* with respect to the set of all real sequences. We now show that Y contains a convergent sequence. Let B denote a rational basis of the real numbers, and let $\{e_p\} = \{a_p b_p\}$ be a null sequence such that $a_i \neq 0$, $a_i \in R$, $b_i \in B$, $i=1, 2, 3, \cdots$, and $b_s = b_t$ if and only if $s=t$. Suppose that

$$(*) \qquad \{e_p\} = c_1 d^{(1)} + c_2 d^{(2)} + \cdots + c_n d^{(n)},$$

where $d^{(i)} \in R'$, $i=1, 2, \cdots, n$, and each c_i is real and nonzero. By the B-representation of a real nonzero number c we will mean the unique (no zero coefficient) representation of c as a finite linear combination with rational coefficients of elements of B. Let N be a positive integer such that the B-representation of c_i does not involve b_N, $i=1, 2, \cdots, n$. Then from (*) we see that b_N is a finite linear combination with rational coefficients of elements of B which are distinct from b_N. But this contradicts the fact that B is a rational basis of the reals. Thus $\{e_p\}$ is a convergent sequence in Y. Let $W = \{X \cup Y, <\}$ be a well-ordering of $X \cup Y$ such that if $x \in X$ and $y \in Y$, then $x < \{e_p\} \leqq y$. Using W, we now form a real basis Z for the set of all real sequences in such a way that $X \subset Z$ and $\{e_p\} \in Z$. Define a sequence-to-sequence transformation A as follows:

(a) If $x \in X$, then $Ax = x$,
(b) A transforms $\{e_p\}$ into $\{1, 0, 1, 0, \cdots\}$,
(c) if $y \in Z - X$ and $y \neq \{e_p\}$, then $Ay = \{0, 0, 0, \cdots\}$,
(d) A transforms $\{0, 0, 0, \cdots\}$ into $\{0, 0, 0, \cdots\}$,
(e) if z is a real sequence, $z \neq \{0, 0, 0, \cdots\}$, and $a_1 z^{(1)} + a_2 z^{(2)} + \cdots + a_q z^{(q)}$ is the Z-representation of z (a_i real, $z^{(i)} \in Z$, no $a_i = 0$), then $Az = a_1 Az^{(1)} + a_2 Az^{(2)} + \cdots + a_q Az^{(q)}$.

Clearly, A is linear and satisfies (1), (2), and (3).

We conclude with a proof of Theorem 1. Suppose (1) holds, but (2) is false. If $\{x_p\} \in M'$ and is convergent, then there exist numbers N and k such that $x_n = k$ if $n > N$. Therefore, if $A = (a_{pq})$ is a matrix such that

(i) $\{a_{np}\}_{n=1}^{\infty}$ converges, $p = 1, 2, 3, \cdots$,

and

(ii) $\{\sum_{p=1}^{\infty} a_{np}\}_{n=1}^{\infty}$ converges,

then A is convergence preserving over M'. Furthermore, if A fails to satisfy

(iii) there exists L such that $\sum_{p=1}^{\infty} |a_{np}| < L$, $n = 1, 2, 3, \cdots$,

then A is not convergence preserving. We note that (i), (ii), and (iii) are the

well-known Silverman-Toeplitz conditions which are necessary and sufficient for A to be convergence preserving. Hence (1) implies (2).

Now suppose that (2) holds and that A is a matrix which is convergence preserving over M'. Let s and t be distinct nonzero numbers in M and let k be a positive integer. Let $z_p=s$ if $p\neq k$, $z_k=t$, and let $y_p=s$, $p=1, 2, 3, \cdots$. Since A transforms $\{z_p\}$, $\{y_p\}$, and consequently $\{z_p-y_p\}$ into convergent sequences, we see that (i) and (ii) hold. Suppose (iii) to be false. Let $A=B+iC$, where B and C are real matrices. Clearly B and C are convergence preserving over M' and satisfy conditions corresponding to (i) and (ii), but one of them must fail to satisfy a condition corresponding to (iii); let us say B. Suppose that r is a limit point of M and let $M_1=\{x-r \mid x\in M\}$ if r is a limit point from the right; otherwise, let $M_1=\{r-x \mid x\in M\}$. Thus 0 is a limit point of M_1 from the right. From the linearity of B and the fact that B satisfies a condition corresponding to (ii), we see that B is convergence preserving over M_1'. Using only slight variations of a method now classical [2, p. 46], we can determine a null sequence in M_1' which B fails to sum. Thus the assumption that (iii) fails to hold is false, and therefore (2) implies (1).

References

1. D. F. Dawson, On certain sequence to sequence transformations which preserve convergence, Proc. Amer. Math. Soc., 14 (1963) 542–545.

2. G. H. Hardy, Divergent series, Clarendon Press, Oxford, 1949.

A NOTE ON THE OUTER AUTOMORPHISMS OF FINITE NILPOTENT GROUPS

O. J. Huval, University of Southwestern Louisiana

Since a finite nilpotent group is a direct product of its Sylow p-subgroups its automorphism group is a direct product of the automorphism groups of its Sylow p-subgroups. To show that a finite nilpotent group has outer automorphisms, then, is to show that one of its Sylow p-subgroups has outer automorphisms, or generally, to show that any finite p-group has outer automorphisms. In this note we give an explicit outer automorphism for certain finite non-abelian p-groups.

In the following material G is a finite non-abelian p-group; $Z(A)$ is the center of group A; $(a, b)=a^{-1}b^{-1}ab$ for any a and b in a group A; $[B, C, \cdots]$ is the set generated by the subsets $B, C, \cdots$ of a group A; $Z_2(A)$ is the complete inverse image of the center of $A/Z(A)$ for any group A; and $\bar{a}$ is the inner automorphism induced by the element a of a group A.

We remark that since $Z_2(A)$ is the complete inverse image of the center of $A/Z(A)$, for any a in A and any b in $Z_2(A)$ (a, b) is in $Z(A)$.

Lemma. *Suppose that the factor group of the normal subgroup M of G is cyclic of order p^q. Let n be such that $[n, M]=G$. If there is an element x in $Z(M)$ so that $(nx)^{p^q}=n^{p^q}$, the mapping ϕ sending $g=n^im$ into $(nx)^im$, where $1\leqq i\leqq p^q$, and $m\in M$ is an automorphism of G under which M is left elementwise fixed.*

Proof. Since the mapping ϕ permutes the elements of each coset and since an element of G has a unique representation modulo M under the conditions in the hypothesis, it follows that

(1) for g and h in G, $g=h$ if and only if $\phi(g)=\phi(h)$.

Since x is in $Z(M)$ and since $(nx)^{p^q}=n^{p^q}$, we see by direct calculation that

(2) for g and h in G, $\phi(gh)=\phi(g)\phi(h)$.

From (1) and (2) it follows that ϕ is an automorphism of G.

Now suppose $m\in M$. Then $m=n^{p^q}m'$ for some m' in M. Then $\phi(m)=(nx)^{p^q}m'$ $=n^{p^q}m'=m$. This shows that M is left elementwise fixed by ϕ.

THEOREM. *Suppose that M and N are maximal subgroups of G. If $Z(M\cap N)\cap Z_2(G)$ properly contains $Z(M)Z(N)\cap Z_2(G)$, then G has outer automorphisms.*

We now give an example of such a group. Let G be the group generated by a, b, and c, where $a^8=b^2=c^2=1$, $bab=a^3$, $cac=a^7$, and $bc=cb$. The subgroup M generated by a and b, and the subgroup N generated by a and c are maximal subgroups of G which satisfy the hypothesis of the theorem.

Proof. We show first that $M\not\subset N$ and $N\not\subset M$. Suppose that $N\subset M$. Then $N\cap M=N$, and $Z(M\cap N)\cap Z_2(G)=Z(N)\cap Z_2(G)\subset Z(M)Z(N)\cap Z_2(G)$. But by hypothesis $Z(M\cap N)\cap Z_2(G)$ properly contains $Z(M)Z(N)\cap Z_2(G)$. Hence, $N\not\subset M$. Similarly, $M\not\subset N$. Since M is maximal in G and since $M\not\supset N$, there is an n in N such that $[n, M]=G$. Similarly there is an m in M such that $[m, N]$ $=G$.

Let x be an element of $Z(M\cap N)\cap Z_2(G)$ not in $Z(M)Z(N)\cap Z_2(G)$. Since $x\in Z(M\cap N)$ and since $n^p\in M\cap N$, $(x^{-1}nx)^p=x^{-1}n^px=n^p$. Thus $(n(n, x))^p=n^p$ for $x^{-1}nx=n(n, x)$. Similarly we show that $(m(m, x))^p=m^p$.

Let g be any element of G. Then, modulo M, $g=n^im'$ for some integer $1\leqq i\leqq p$ and some m' in M and, modulo N, $g=m^jn'$ for some integer $1\leqq j\leqq p$ and some n' in N. Since $(n, x)\in Z(M)$ and $(n(n, x))^p=n^p$, it follows from the above lemma that the mapping α sending $g=n^im'$ into $(n(n, x))^im'$ is an automorphism of G under which M is left elementwise fixed. Similarly the mapping β sending $g=m^jn'$ into $(m(m, x))^jn'$ is an automorphism of G under which N is left elementwise fixed.

From the definitions of α and β it follows that $\alpha\beta=\bar{x}$. Since $\bar{x}$ is an inner automorphism, then α and β are both inner automorphisms, or they are both outer automorphisms. Suppose that α and β are both inner. Then there exist u in $Z(M)$ and v in $Z(N)$ so that $\alpha=\bar{u}$ and $\beta=\bar{v}$. Since $\alpha\beta=\bar{x}$, $\bar{x}=\bar{u}\bar{v}$. But this implies that x is in $Z(M)Z(N)\cap Z_2(G)$. Hence α and β are not inner automorphisms but outer automorphisms.

Reference

1. H. J. Zassenhaus, The Theory of Groups, Chelsea, New York, 1958.

there is $a \in A$ such that $ax \neq 0$. $Ra \cap x^l \neq (0)$ since x^l is large. Hence there is $r_1 a \neq 0$ in Ra such that $(r_1 a)x = 0$. This is impossible since $(r_1 a)^r = a^r$.

THEOREM 3.2. *Let $R \in C$ and $L_r^*(R)$ be atomic. If R has a regular ring Q as a two-sided quotient ring* (in the sense of R. E. Johnson) *then $L_l^*(R)$ is also atomic. If, in addition, some uniform right ideal of R is a prime ring, then Q is a division ring and R is a right and left Ore-domain.*

Proof. If Q is a regular ring then each atom of $L_r^*(Q)$ is a minimal right ideal by Lemma 2.3. Let $\overline{A}$ be an atom of $L_r^*(Q)$. Then $\overline{A} = eQ$ for some idempotent e in Q. Since $\overline{A}$ is a minimal right ideal, Qe is a minimal left ideal. Thus $Qe \cap R$ is an atom of $L^*(R)$. In case atom $A \in L_r^*(R)$ is a prime ring then $(\overline{A})^l = (0)$ by Lemma 2.2. Hence e must be equal to 1. Thus $\overline{A} = Q$ and Q is a division ring.

REMARK. The Theorem 3.2 asserts that in general the maximal two-sided quotient ring of a ring R with $R_r^\Delta = (0)$ and $R_l^\Delta = (0)$ is not necessarily regular.

Acknowledgment. I am deeply indebted to the referee for many helpful suggestions.

References

1. R. E. Johnson, The extended centralizer of a ring over a module, Proc. Amer. Math. Soc., 2 (1951) 891–895.

2. ———, Structure theory of faithful rings II. Restricted rings, Trans. Amer. Math. Soc., 84 (1957) 523–544.

3. ———, Quotient rings of rings with zero-singular ideal, Pacific J. Math., Vol. II, No. 4 (1961) 1385–1392.

A TYPE OF "GAMBLER'S RUIN" PROBLEM

R. C. READ, University of the West Indies, Jamaica

In his book "Tomorrow's Math" ([1], page 114) Ogilvy quotes the following as an unsolved problem: "Three men have respectively l, m and n coins which they match so that the odd man wins. Whenever all coins appear alike they repeat the throw. Find the average number of tosses required until one man is forced out of the game." In this paper we derive the solution to this problem, which turns out to be unexpectedly simple.

Let $E(l, m, n)$ be the expected number of tosses. Of course we assume that the coins are "fair". If one toss is made then there is

(i) probability $\frac{1}{4}$ that the coins appear alike, in which case $E(l, m, n)$ further tosses are expected.

(ii) probability $\frac{1}{4}$ that the first player wins, in which case $E(l+2, m-1, n-1)$ further tosses are expected.

(iii) and (iv) probabilities as for (ii) but for the second or third players.

From these remarks we see that

$$E(l, m, n) = \tfrac{1}{4}[1 + E(l, m, n)] + \tfrac{1}{4}[1 + E(l + 2, m - 1, n - 1)]$$
$$+ \tfrac{1}{4}[1 + E(l - 1, m + 2, n - 1)] + \tfrac{1}{4}[1 + E(l - 1, m - 1, n + 2)]$$

or

$$\tfrac{3}{4}E(l, m, n) = 1 + \tfrac{1}{4}E(l + 2, m - 1, n - 1) + \tfrac{1}{4}E(l - 1, m + 2, n - 1) + \tfrac{1}{4}E(l - 1, m - 1, n + 2).$$

It will be convenient to put $l=p+1$, $m=q+1$, $n=r+1$ and thus obtain

$$(1) \qquad E(p + 1, q + 1, r + 1) = \tfrac{4}{3} + \tfrac{1}{3}E(p + 3, q, r) + \tfrac{1}{3}E(p, q + 3, r) + \tfrac{1}{3}E(p, q, r + 3).$$

Throughout the game the sum $l+m+n=p+q+r+3$, the total number of coins, remains the same. Therefore although equation (1) has the appearance of a recursion formula, it cannot be used directly to calculate $E(l, m, n)$ in terms of E's with smaller values of l, m and n, as one would expect from a recursion formula. For example, if $l+m+n=7$ then, since $E(l, m, n)=0$ if any one of l, m or n is zero, we can derive from (1) the following equations:

$$E(5, 1, 1) = E(1, 5, 1) = E(1, 1, 5) = \tfrac{4}{3}$$
$$E(4, 2, 1) = \tfrac{4}{3} + \tfrac{1}{3}E(3, 1, 3)$$
$$E(3, 3, 1) = \tfrac{4}{3} + \tfrac{1}{3}E(2, 2, 3)$$
$$E(3, 2, 2) = \tfrac{4}{3} + \tfrac{1}{3}E(5, 1, 1) + \tfrac{1}{3}E(2, 4, 1) + \tfrac{1}{3}E(2, 1, 4).$$

These three equations can be solved and we find, since the function E is clearly symmetrical in its arguments, that

$$E(4, 2, 1) = E(2, 4, 1) = E(2, 1, 4) = \tfrac{32}{15}$$
$$E(3, 3, 1) = E(3, 1, 3) = E(1, 3, 3) = \tfrac{12}{5}$$

and

$$E(2, 2, 3) = E(2, 3, 2) = E(3, 2, 2) = \tfrac{16}{5} \cdot$$

This method of calculating the values of $E(l, m, n)$ will clearly become increasingly tedious as the value of $l+m+n$ increases, and in any case is not likely to yield a general formula. Moreover the usual techniques of using counting series or generating functions do not seem to be of much help in a problem of this kind.

Let us turn to a simpler problem. Two men match coins. If one is "heads" and the other "tails," then "heads" wins. If the coins are alike, neither player wins. As before, we must find the expected number of tosses before one player is ruined. It is easily seen that this problem is equivalent to a random walk on a line, with probabilities $\frac{1}{4}$, $\frac{1}{2}$ and $\frac{1}{4}$ of moving one unit to the left, staying put, or moving one unit to the right. The equation analogous to (1) is

$$(2) \qquad E(p + 1, q + 1) = 2 + \tfrac{1}{2}E(p + 2, q) + \tfrac{1}{2}E(p, q + 2)$$

with an obvious notation. If we write $u_l=E(l, N-l)$, where N is the total number of coins, then (2) becomes

$$u_{p+2} - 2u_{p+1} + u_p = -4. \tag{3}$$

This is a straightforward difference equation, the solution of which, by the usual methods is

$$u_p = Cp + D - 2p(p-1),$$

where C and D are constants. We require u_p to be zero when $p=0$ and when $p=N$. For this we must have $D=0$ and $C=2(N-1)$, and we obtain

$$E(l, m) = 2(N-1)l - 2l(l-1) = 2lm.$$

This is an extremely simple result, and it would be sanguine in the extreme to expect a similar result to hold for equation (1), but there is no harm in trying. Let us assume $E(l, m, n) = Klmn$, where K is a constant. Equation (1) becomes

$$K(p+1)(q+1)(r+1) = \tfrac{4}{3} + \tfrac{1}{3}K\{(p+3)qr + p(q+3)r + pq(r+3)\} \tag{4}$$

or $K(p+q+r)+K=\frac{4}{3}$, since the terms of the second and third degrees cancel. Hence

$$K = \frac{4}{3(p+q+r+1)} = \frac{4}{3(l+m+n-2)}, \text{ a constant!}$$

Thus we have $E(l, m, n) = (4lmn)/3(l+m+n-2)$, which furnishes one solution of the problem much more easily than one would have any reason to expect.

A similar problem with four players would be unlikely to have a simple solution. In the equation analogous to (4), in fact, the terms of degree 4 would cancel, those of degree 3 might do so, but those of degree 2 would almost certainly not. Hence no constant K would cause the equation to be satisfied. A typical problem of this type would be the following. Four players match coins, the odd man winning and collecting a coin from each of the other three. If all coins are alike, or if there are two heads and two tails, then no coins change hands. Given the number of coins each player has in the beginning, find the expected number of tosses before one player loses all his coins. Other problems with four players will result from different rulings concerning the way the coins change hands.

It is readily shown that the function $Kl_1l_2l_3l_4$, where l_1, l_2, l_3, l_4 are the numbers of coins, is *not* the solution to this problem. Other plausible possibilities, such as a linear combination of the symmetric functions $l_1l_2l_3l_4$, $\Sigma l_1l_2l_3$, Σl_1l_2 and Σl_1 can also be easily ruled out. It looks as though this problem is one that does not have an easy solution.

Reference

1. C. Stanley Ogilvy. Tomorrow's Math; unsolved problems for the amateur, Oxford University Press, New York, 1963.

Editorial Note. This is Problem 4003 in this MONTHLY, 48 (1941) 483, proposed by G. W. Petrie. The galley proofs of the present issue contained a solution by F. Göbel (Netherlands).

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

"THE PRICE IS RIGHT" GAME

SHIMON EVEN, Sperry Rand Research Center, Sudbury, Mass.

Let there be n players $A_1, A_2, \cdots, A_n$ playing on the interval $[0, 1]$ in the following way: Each player in his turn (A_1 plays first, and they follow in the natural order in a cyclic manner) either makes a bid $0 \leqq x \leqq 1$ such that x is larger than any previously made bid, or passes, in which case his last bid becomes his final bid. The payoff for each player is computed after the bidding is over, and is defined by the length of the interval between his final bid x and the least bid higher than x or, in the absence of such a bid, it is $1-x$. The payoff for a player who makes no bid is zero, and therefore, each player will make at least one bid. (This game was inspired by the ABC television show of the same name. The rules have been changed, however, and the assumption has been made that all the players have the same price distribution in mind.)

It is clear that in this game the players are not really playing against their opponents (this is not a zero-sum game), but rather do their best to maximize their own payoffs subject to the knowledge that their opponents have the same objective in mind. It is assumed that the players are rational, and are capable of contemplating all possible outcomes. Also it is assumed that if by bidding again a player can get at most as much as by passing, he will choose to pass.

The 2-player game. For the case $n=2$ it will be shown that there exists a unique way of playing the game. Surprisingly, the solution is given by the famous golden section.

Player A_1 will start the bidding by naming 0 as his first bid, because if he bids $1 \geqq x > 0$ then the game will proceed on the interval $[x, 1]$ in the same manner as it would on $[0, 1]$, only that now the final payoffs will be smaller by a factor $1-x$. The second player's first bid will be called *terminal*, if the first player will respond with a "pass" and thus terminate the bidding. It is clear that there exist terminal moves. For example $\frac{1}{2}$ is terminal, because after A_2 has bid $\frac{1}{2}$, A_1 can name only $1 \geqq x > \frac{1}{2}$, by which his final payoff is sure to be less than $\frac{1}{2}$, and if he passed his payoff is $\frac{1}{2}$. It is easy to see that there also exist terminal moves less than $\frac{1}{2}$. Let λ be the least terminal bid.

THEOREM. *In the 2-player game, A_2 will bid λ on his first bid.*

Proof. Clearly, A_2 will not bid $1 \geqq x > \lambda$, since x is also terminal, and A_2's

payoff is thus reduced. Assume now that A_2 plays $\lambda > x \geqq 0$. Since λ is the least terminal bid, A_1 will not pass. By bidding $x+\lambda(1-x)$, A_1 terminates the bidding and assures himself a payoff of $(1-\lambda)(1-x)$. Therefore he will bid in some way that will produce for him at least $(1-\lambda)(1-x)$. Clearly, then, A_2 cannot expect to get a payoff higher than $(1-x)-(1-\lambda)(1-x)=\lambda(1-x)$, and since $\lambda<\frac{1}{2}$, and $(1-x)<1$, he will get a payoff less than $\frac{1}{2}$, whereas by playing λ he assures himself a payoff greater than $\frac{1}{2}$.

Now that it has been established that λ is the best strategy for the second player, it is possible to compute λ. For if A_1 decides to bid again, after A_2 has bid λ, the best he can do is bid $\lambda+\lambda(1-\lambda)$, which will terminate the bidding, and his new payoff is $(1-\lambda)^2$. In order to be indifferent about this move, so that λ is indeed terminal, it is required that

$$\lambda \geqq (1-\lambda)^2,$$

and the minimal value of λ for which this holds is given by: $\lambda=(1-\lambda)^2$. The solution of this equation (in $[0, 1]$) is

$$\lambda = \frac{3-\sqrt{5}}{2},$$

which is the well-known value of the golden section. In short the 2-player game will proceed in the following way: (a) A_1 will bid 0; (b) A_2 will bid $\lambda=(3-\sqrt{5})/2$; (c) A_1 will pass; (d) A_2 will pass.

Clearly, A_2 has the advantage, and he will get a payoff of $(1-\lambda)\cong 0.618$, while A_1 gets $\lambda\cong 0.382$.

The n-player game. The author has found the n-player game, for $n\geqq 3$, hard to analyze. He will report in this section his progress and conjectures.

In the 3-player game, if any two players form a coalition they can keep their opponent's payoff below any $\epsilon>0$. The n-player game, for $n\geqq 4$, with coalitions is extremely difficult to analyze, and no progress has been made in this direction. Throughout the rest of this section it will be assumed that *no coalitions are allowed.*

Consider the following sequence of bids for the 3-player game:

(a) A_1 bids 0, (b) A_2 bids $\dfrac{\lambda}{1+\lambda}$.

Faced with this situation, A_3 will bid $2\lambda/1+\lambda$. To demonstrate this fact, it will be shown first that this terminates the bidding.

If A_1 passes, A_2 passes too, for he is now in the same situation as the first player is in the 2-player game after the second player bid λ. (This follows from the fact that $[\lambda/(1+\lambda)]/[1-\lambda/(1+\lambda)]=\lambda$.) Also, it is clear that A_1 will pass, for he can get at most what he is already assured of having. Thus, if A_3 bids $2\lambda/1+\lambda$ he terminates the bidding. If he bids $x>(2\lambda/1+\lambda)$, he still terminates the bidding and has less payoff. If he bids $x<(2\lambda/1+\lambda)$, and A_1 passes, A_2 is

sure to bid again and improve his payoff, for

$$(1-\lambda)(1-x) > (1-\lambda)\left(1-\frac{2\lambda}{1+\lambda}\right) = \frac{(1-\lambda)^2}{1+\lambda} = \frac{\lambda}{1+\lambda};$$

while if A_1 chooses to move, he is assured of getting this much himself. In either case the payoff to A_3 will be decreased. Thus, if A_2 bids $\lambda/1+\lambda$ then A_3 will follow with $2\lambda/1+\lambda$ and terminate the bidding. The natural question to ask is: can A_2 improve his payoff by bidding $x<(\lambda/1+\lambda)$? (Clearly, $x>(\lambda/1+\lambda)$ will not do.) If A_3 still does best by terminating the bidding, he has to move to the right, thus widening A_2's payoff even further. It is conjectured that A_3 will choose that strategy as long as he can not get more by the same scheme on the interval $1-x$. The value of x can now be evaluated by solving the equation

$$(1-x)\left(1-x-\frac{x}{1-\lambda}\right) = \frac{x}{1-\lambda}\,.$$

The solution is given by $\mu=1-\sqrt{(1-\lambda)}$. Thus, it is conjectured that the 3-player game will proceed as follows: (a) A_1 bids 0; (b) A_2 bids $\mu\cong.214$; (c) A_3 bids $1-(\mu/1-\lambda)\cong.654$; (d) Everyone passes. This means that the second player is the most powerful, and his payoff is about 44 percent of the whole interval.

It is conjectured that the n-player game proceeds in a similar manner; namely, that each player bids only once. Let $a_1^n, a_2^n, \cdots, a_n^n$ be the bids made by $A_1, A_2, \cdots, A_n$. Clearly, $a_1^n=0$. It is conjectured that

$$\begin{aligned}
a_2^{n+1} &= a_2^n(1-a_3^{n+1}+a_2^{n+1})\\
a_4^{n+1}-a_3^{n+1} &= (a_3^{n+1}-a_2^{n+1})(1-a_2^{n+1})\\
a_5^{n+1}-a_4^{n+1} &= (a_4^n-a_3^n)(1-a_3^{n+1}+a_2^{n+1}) \qquad \text{(define } a_4^3=1)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
a_{n+1}^{n+1}-a_n^{n+1} &= (a_n^n-a_{n-1}^n)(1-a_3^{n+1}+a_2^{n+1})\\
1-a_{n+1}^{n+1} &= (1-a_n^n)(1-a_3^{n+1}+a_2^{n+1})
\end{aligned}$$

describe the solution for the $(n+1)$-player game, where for $n=2$ the first two equations are used instead of the first and the last.

Acknowledgment. The author wishes to thank Drs. I. M. Isaacs and W. L. Eastman and Mr. A. A. Hauser, Jr. for stimulating discussions.

CONVERGENCE ALMOST EVERYWHERE IS NOT TOPOLOGICAL

E. T. ORDMAN, Princeton University

When the notion of convergence almost everywhere (a.e.) is introduced, the student acquainted with convergence in general topology may try to carry over more properties than are justified. This note shows that if X is the space of bounded measurable real functions on [0, 1], convergence a.e. is not equivalent to convergence in any topology on X.

Let $\{f_n\}$ be a sequence of functions which converges in measure to zero but fails to converge a.e.; the standard example is the sequence $f_1^1, f_1^2, f_2^2, f_1^3, f_2^3, \cdots$, where

$$f_m^n(x) = \begin{cases} 1 & (m-1)/n \leqq x \leqq m/n \\ 0 & \text{otherwise} \end{cases} \qquad 1 \leqq m \leqq n.$$

Now suppose that a topology of convergence a.e. exists on X. Since $\{f_n\}$ fails to converge to zero, there must be a neighborhood $N(0)$ which f_n is frequently outside; let $\{f_{n'}\}$ be the subsequence of terms outside of $N(0)$. Then $\{f_{n'}\}$ converges in measure to zero, so by a standard theorem ([1], p. 46), it has a subsequence which converges a.e. to zero. But that subsequence is eventually in $N(0)$, contradicting the choice of $\{f_{n'}\}$ to remain outside. That is, the supposed topology cannot exist; in fact, there is no way to define convergence a.e. by a neighborhood filter of the usual sort.

Reference

1. A. N. Kolmogorov and S. V. Fomin, Functional Analysis, Vol. 2, Graylock, Albany, 1961.

AN ELEMENTARY CONSTRUCTION OF A FINITE NONARGUESIAN PROJECTIVE PLANE

J. R. WESSON, Vanderbilt University

1. Introduction. We give here a construction of a known finite nonarguesian plane in such a way that beginning students can easily check the details. Our construction is equivalent to one that is well-known [1, p. 408; 2, p. 364].

In Section 2, a Veblen-Wedderburn system with 9 elements is presented. The corresponding projective plane is given in section 3. Also, we outline six cases in which the theorem of Desargues fails.

The paper can be read easily by a student with an elementary course in abstract algebra, and the author believes the material could be used in connection with an elementary course in algebra or geometry.

2. Construction of a Veblen-Wedderburn system. Let F be the field of nine elements, with the usual identities 0, 1, and with $2=1+1$. (Recall that F can be displayed as the set of elements $m+ni$, where m, n represent residue classes modulo 3. Addition is given by $(m_1+n_1i)+(m_2+n_2i)=(m_1+m_2)+(n_1+n_2)i$, and multiplication makes use of the definition that $i^2+1=0$ [1, p. 410].)

An element x of F is called *square* provided $x=a^2$ for some $a \in F$. It is easy to prove the following for $x \in F$:

(2.1) $2x^5+2x=x$ or 0, according to whether x is square or not square. (The square elements are 0, 1, 2, i, and $2i$.)

(2.2) $x^5+2x=0$ or x, according to whether x is square or not square.

(2.3) $x^2+x^4+x^6=2$ for all x except 0, 1, 2.

Next we use the elements of F to form a Veblen-Wedderburn system V.

Addition in V is the same as in F, and we define a new multiplication $\cdot$ for V by

$$x\cdot y = x(2y^5 + 2y) + x^3(y^5 + 2y), \tag{2.4}$$

where operations in the right member are from F. From (2.1), (2.2), this is equivalent to defining $x\cdot y = xy$ for y square (*a square element of F*), and $x\cdot y = x^3y$ for y not square [1, p. 410].

In order to see that V is a Veblen-Wedderburn system, we check the following properties.

P1. $0\cdot x = x\cdot 0 = 0$; $1\cdot x = x\cdot 1 = x$.

P2. *If $a, b \in V$ with $a \neq 0$, then the equation $x\cdot a = b$ has a unique solution x, and $a\cdot y = b$ has a unique solution y.*

Proof. If a is square, $x\cdot a = xa$. If not, $x\cdot a = b$ if and only if $x^3a = b$. Hence $x\cdot a = b$ is uniquely solvable.

For the remainder, it is sufficient (since V is finite) to prove that $a\cdot y_1 = a\cdot y_2$ implies $y_1 = y_2$. For the only nontrivial case, let y_1 be square and y_2 not square. Then $ay_1 = a^3y_2$, $y_1 = a^2y_2$, and y_2 is square, a contradiction.

P3. $(x+y)\cdot z = x\cdot z + y\cdot z$.

Proof. If z is not square, use the identity $(x+y)^3 = x^3 + y^3$ from F.

P4. *If $a, b, c \in V$ with $a \neq b$, then $x\cdot a = x\cdot b + c$ has a unique solution x.*

Proof. The proof is immediate if a, b are either both square or both not square. Suppose a is square and b is not. Then $xa = x^3b + c$ if and only if $x^3b - xa + c = 0$. To show that this last equation has exactly one root, it suffices to show that $x_1^3 - x_1ab^{-1} + cb^{-1} = x_2^3 - x_2ab^{-1} + cb^{-1}$ implies $x_1 = x_2$. If $x_1^3 - x_2^3 = x_1ab^{-1} - x_2ab^{-1}$ with $x_1 \neq x_2$, then $(x_1 - x_2)^2 = ab^{-1}$ and b is square, a contradiction.

We can cover the case for b square and a not square by applying the preceding to $x\cdot b = x\cdot a - c$.

LEMMA 1. *Let a, b, c, d be any elements with $a \neq c$. The system $a\cdot x + y = b$, $c\cdot x + y = d$ has a unique solution (x, y).*

Proof. If $a = 0$, the solution is given by $y = b$, $c\cdot x = d - b$. If $a \neq 0$, $a\cdot x_i + y_i = b$ determines a 1-1 correspondence $x_i \leftrightarrow y_i$. Also, if $x_i \neq x_j$, then $c\cdot x_i + y_i \neq c\cdot x_j + y_j$, for otherwise the pair $c\cdot x_i + y_i = c\cdot x_j + y_j$, $a\cdot x_i + y_i = a\cdot x_j + y_j$ would contradict the uniqueness mentioned in P4. Hence $c\cdot x_i + y_i = d$ for exactly one pair x_i, y_i.

LEMMA 2. *If $x \neq 0, 1, 2$, then $x\cdot x = 2$.*

Proof. From (2.4) and (2.3), $x\cdot x = 2(x^2 + x^4 + x^6) + x^8 = 2$.

LEMMA 3. $2\cdot x = 2x$.

Proof. $2\cdot x = (1+1)\cdot x = x + x = 2x$.

LEMMA 4. $2\cdot(x+y)=2\cdot x+2\cdot y=2x+2y$.

Proof. Use Lemma 3.

3. Construction of the nonarguesian projective plane. We use the construction in Hall [2, p. 353]. Points are taken as (a, b), (m), (∞), where a, b, m are any elements of V, and $\infty \notin V$. Lines are L_∞, $x=a$, $y=x\cdot m+b$, where $a, m, b\in V$. The following (and no other) incidences between points and lines are defined. The line $x=a$ contains (∞) and all points (a, b). The line L_∞ contains (∞) and all points (m). The line $y=x\cdot m+b$ contains (m) and all points (x_1, y_1) such that $y_1=x_1\cdot m+b$.

THEOREM. *Let b be any element of V other than* 0, 1, 2. *The triangles $A(1, 1)$, $B(2b+1, b+2)$, $C(1, b)$ and $A'(2b+1, 1)$, $B'(2, 2b+1)$, $C'(2, 2b+2)$ are centrally perspective, but Desargues' theorem fails.*

Proof. Line AA' has equation $y=1$. The equation of BB' is $y=x\cdot b+1$, for

$$(2b+1)\cdot b+1=(b+b+1)\cdot b+1=b\cdot b+b\cdot b+b+1=b+2$$

by P3 and Lemma 2, and $2\cdot b+1=2b+1$ by Lemma 3. Line CC' has equation $y=x\cdot(b+2)+1$, for $1\cdot(b+2)+1=b$, and $2\cdot(b+2)+1=2b+2$ (Lemma 4). Hence triangles ABC and $A'B'C'$ are perspective from (0, 1).

The following lines and points have equations or coordinates as indicated.

BC:	$y=x\cdot(2b)+2b$	$B'C'$:	$x=2$
CA:	$x=1$	$C'A'$:	$y=x\cdot(2b)+b+2$
AB:	$y=x\cdot(2b+1)+b$	$A'B'$:	$y=x\cdot(b+2)$

D or $(BC, B'C')$, the intersection of BC, $B'C'$: (2, 0)
E or $(CA, C'A')$: (1, 2)
F or $(AB, A'B')$: $(b+1, 2b)$
DE: $y=x+1$.

For example, B lies on $y=x\cdot(2b)+2b$, since

$$\begin{aligned}(2b+1)\cdot(2b)+2b&=(2b)\cdot(2b)+2b+2b\\&=2+b,\qquad\text{(apply Lemma 2 for } x=2b).\end{aligned}$$

Similarly, B lies on $y=x\cdot(2b+1)+b$, since if $b\neq 0, 1, 2$ then $2b+1\neq 0, 1, 2$, and, by Lemma 2, $(2b+1)\cdot(2b+1)+b=2+b$.

Also, A' lies on $y=x\cdot(b+2)$, since

$$\begin{aligned}(2b+1)\cdot(b+2)&=(b+2+b+2)\cdot(b+2)=(b+2)\cdot(b+2)+(b+2)\cdot(b+2)\\&=2+2=1,\qquad\text{(use Lemma 2).}\end{aligned}$$

In the same way, $(b+1, 2b)$ lies on $y=x\cdot(2b+1)+b$, for $(b+1)\cdot(2b+1)+b$

$=b\cdot(2b+1)+1=(2b+1+2b+2)\cdot(2b+1)+1=2+(2b+2)\cdot(2b+1)+1=(2b+1)\cdot(2b+1)+(2b+1)=2+2b+1=2b$, (use P3 and Lemma 2).

Similarly, $(b+1,\ 2b)$ lies on $y=x\cdot(b+2)$, since $(b+1)\cdot(b+2)=(b+2+2)\cdot(b+2)=2+2\cdot(b+2)=2b$.

The other incidences are easier to check. Since F does not lie on line DE, the plane is nonarguesian.

References

1. R. D. Carmichael, Groups of finite order, Ginn, New York, 1937.
2. M. Hall, The theory of groups, Macmillan, New York, 1959.

A CLASS OF NONISOMORPHIC RINGS, EACH ISOMORPHIC TO A PROPER SUBRING OF EVERY OTHER

B. R. Toskey, Seattle University

We give here an example of a countably infinite collection of commutative rings, all having the same additive group, having the properties 1) no two are isomorphic to each other, and 2) any ring in the collection will possess an infinite number of subrings which are isomorphic to any other selected ring.

The additive group common to all rings in the class we will consider is the direct sum of two infinite cyclic groups, with generators u and v. The rings we construct will be denoted by T_n, $n=1, 2, 3, \cdots$, and we need only to define multiplication in T_n. In the ring T_n, we shall require:

$$\begin{cases} uu = nv \\ uv = vu = vv = 0 \end{cases}$$

and extend by the distributive law to all members of T_n. Since all triple products are zero, associativity is trivially satisfied, and hence the above products define a commutative ring.

A. *If T_n and T_m are isomorphic, then $n=m$.*

Proof. Suppose we have an isomorphism α: $T_n \cong T_m$ so that

$$\alpha(u) = au + bv, \qquad \alpha(v) = cu + dv,$$

for some integers a, b, c, d, where the images of α are in T_m. By the homomorphism property, we must have:

$$\alpha(uu) = \alpha(nv) = a^2mv = ncu + ndv$$
$$\alpha(vv) = \alpha(0) = c^2mv = 0,$$

from which we obtain $c=0$ and $nd=a^2m$. Since α is onto:

$$\alpha(eu + fv) = u \qquad \alpha(gu + hv) = v$$

for some integers e, f, g, h, and we obtain $ea=1$, $g=0$, and $hd=1$. This implies that $a^2=d^2=1$ and hence $n=m$, since both are positive.

B. *For any n, m, T_m has an infinite number of subrings which are isomorphic to T_n.*

Proof. Choose any integer $r \neq 0$ and any integer s, and let: $\beta(u) = nru + sv$, $\beta(v) = r^2 nmv$ define an additive mapping on T_n into T_m, extended by linearity to all of T_n. It is easy to verify that multiplication is also preserved, and hence β is a ring homomorphism. If $au + bv$ is in the kernel of β, we must have $anr = 0$ and $as + br^2 nm = 0$, and since $rnm \neq 0$, we have $a = b = 0$, so that β is an isomorphism. If r is given the values $1, 2, 3, \cdots$ the images are distinct, so that T_m has an infinite number of subrings isomorphic to T_n.

THE NATURAL LOGARITHM IS TRANSCENDENTAL

H. R. ROUSE, Tufts University

If a function f is algebraic, then for some positive integer n, there exists a sequence $P_0, \cdots, P_n$ of polynomial functions such that for every number x in the domain of f

$$\sum_{k=0}^{n} P_k(x)[f(x)]^k = 0. \tag{1}$$

For such a function there is a *least* positive integer n for which such a sequence exists, and for this n there is a *least* positive integer m for which such a sequence exists consisting of polynomials of degree $\leqq m$. Call a sequence with this pair n and m a *minimal* sequence for the function.

We shall prove that there is no algebraic function whose derivative exists and is $1/x$ for every number x in its domain.

Suppose there were such a function f. Clearly 0 is not in the domain of f. Let $P_0, \cdots, P_n$ be a minimal sequence of polynomials for f, and differentiate (1) to obtain

$$\sum_{k=0}^{n} P_k'(x)[f(x)]^k + \sum_{k=1}^{n} kP_k(x)[f(x)]^{k-1}f'(x) = 0. \tag{2}$$

In the second summation in (2) replace $f'(x)$ by $1/x$ and change the index k to $k+1$, getting

$$\sum_{k=0}^{n} P_k'(x)[f(x)]^k + \sum_{k=0}^{n-1} (k+1)P_{k+1}(x)[f(x)]^k \frac{1}{x} = 0. \tag{3}$$

Multiply (3) by x and regroup, obtaining

$$\sum_{k=0}^{n-1} [xP_k'(x) + (k+1)P_{k+1}(x)][f(x)]^k + xP_n'(x)[f(x)]^n = 0. \tag{4}$$

From (4) it follows that there is a sequence of polynomial functions having the same n and m such that the last polynomial has a factor of x. If we assume,

without loss of generality, that $P_n(x)$ has a factor of x, it follows further from (4) that there is such a sequence of polynomial functions whose last *two* polynomials both have a factor of x. Repeating this argument n times, we see that we may assume that each of the polynomials $P_0, \cdots, P_n$ has a factor of x. Now since 0 is not in the domain of f, (1) can be divided by x, reducing the degree of each polynomial. But this contradicts the stipulation that m be the least such integer, and the proof is complete.

An algebraic function for which $n=1$ is a rational function. A good exercise for the student is to follow the pattern of the above proof and show that there is no rational function whose derivative is $1/x$. The better student might be challenged to prove in a similar manner that there is no algebraic function whose derivative exists and equals the function itself throughout its domain. (Actually there is *one* such algebraic function, but he should discover this for himself.)

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College, and
JOHN A. BROWN, University of Delaware

All material for this department should be sent to John R. Mayor, 1515 Massachusetts Avenue, N.W., Washington D. C. 20005

THE RELATIONSHIP BETWEEN THE PERFORMANCE OF THE TEXAS SOUTHERN UNIVERSITY FRESHMEN ON THE MATHEMATICS PLACEMENT TEST AND THEIR HIGH SCHOOL MATHEMATICS BACKGROUND

L. K. BRADLEY, C. R. NEWELL, AND V. M. WILLIAMS, Texas Southern University

Summary. The general purpose of this study was to determine the relationship between the performance of the Texas Southern University freshmen on the mathematics placement test and their high school mathematics background.

At Texas Southern University, the test required for entrance is not used to restrict enrollment. Hence, no minimum mathematical achievement can be assumed. A hasty examination of the records will indicate that an overwhelming majority of the students admitted to Texas Southern University appear to have an adequate mathematical background; yet more than 93% fail to pass the mathematics entrance test.

In developing this problem the following questions were formulated:

A. What is the mathematical background of the students who took the test?
B. To what extent is the mathematical information obtained in high school by these students reflected on the test?
C. What inferences can be drawn from the data?

D. What suggestions may be offered for improvement of the entering students on the mathematics placement test?

The data were obtained by (1) referring to the students' high school transcripts on file in the Registrar's office; (2) analyzing the results of the mathematics placement test; and (3) interviewing a sample of the students involved in the study. The students used in this study were those 806 students (1) who took the mathematics placement test during September, 1962; and (2) whose high school transcripts were on file in the Registrar's office.

Background information, as recorded in the files of the Registrar's office showed that the 806 students had graduated from 204 high schools located in 21 states of the United States and one foreign state. A majority of the high schools (137) were located in Texas with approximately 90% of the students involved in this study being graduates of these high schools. The size of the high school graduating classes ranged from 4 to 716, with the size of the majority of the graduating classes being less than 200. Of the 665 students whose transcripts showed quartile information, 162 or 20.09% of the entire group, ranked in the first quartile of their high school graduating classes.

It was found that 660, or approximately 82% of the total group of students, presented more than two units of high school mathematics. One hundred twenty-one students presented four units of high school mathematics; two students five units; and one student presented zero units of high school mathematics. Mathematics courses taken in high school include, in the order of frequency: algebra I, plane geometry, algebra II, general mathematics, commercial mathematics, trigonometry, solid geometry, business mathematics, advanced mathematics, and basic mathematics. The most frequent combination of high school mathematics courses presented was algebra I, algebra II, and plane geometry.

Two hundred four students, or approximately 25% of the total group of students, achieved an average of "B" or better in their high school mathematics courses; 459, or approximately 57%, obtained an average of "C" and above but less than "B"; and 143, or approximately 18%, achieved an average less than "C."

Information recorded in the Registrar's office indicated that the students were interested in majoring in education, the fine arts, home economics, the humanities, natural and physical sciences, physical and health education, social sciences, business, and vocational and industrial education, with the largest percentage, 20.8%, interested in majoring in the natural and physical sciences.

It was found that on the mathematics placement test, 72% of the students (581) answered correctly one-half or less of the 74 items on the test. Further, the students performed best on the items classified under "Number and Operation" and poorest on the group "Proof-deductive and Inferential Reasoning."

Of the 81 students, chosen at random for interviews, two had passed the test; three did not take the test; and 76 had failed the test. "Inadequate background" was the most frequent response given to the question asked concerning

reasons for failure. Other responses given by the students to the question, in order of frequency, were: (1) thought they had passed, (2) fatigue and nervous tension, (3) had forgotten, (4) insufficient time, and (5) insufficient understanding of marking the answer sheet.

A formula for computing a Pearson product-moment coefficient of correlation was used to determine the relationship between the high school mathematics grades and the scores on the mathematics placement test. The coefficient of correlation was calculated to be .39.

Recommendations. On the basis of the information compiled in this study, the following recommendations are made:

A. *Recommendations Concerning the Students.* From the evidence available, it appears that an overwhelming majority of the students tested have taken enough high school mathematics courses to be well prepared in mathematics, yet the test results indicate that these students have not fully grasped the mathematical material to which they have been exposed. Therefore, it is recommended that:

1. Students be advised that low level work (as evidenced by grades of C and below) does not provide an adequate background for success in college work.

2. Many opportunities be provided for increasing and broadening the mathematical scope of the students.

3. Some formal mathematical training be a part of each student's work during his senior year in high school.

4. All high school students have some experience with algebraic processes before graduation.

5. Classroom situations be so structured as to provide extensive involvement in critical thinking, inasmuch as the students performed poorest on the group of test items classified under "Proof-deductive and Inferential Reasoning."

6. The mathematics staff at Texas Southern University be more diligent in making high school faculties more cognizant of the mathematical difficulties encountered by students entering college.

B. *Recommendations for Further Study.* This study has called attention to certain problems that would bear further investigation. They are:

1. A study to determine the relationship between the grades earned in first year college mathematics and grades earned in the high school mathematics courses.

2. A study designed to determine the relationship between the student's performance on the mathematics placement test and their choices of fields of collegiate major concentration.

This research was supported by a grant from the Faculty Research Committee of Texas Southern University.

COLLEGE INSTRUCTORS OF MATHEMATICS FOR ELEMENTARY-SCHOOL TEACHERS

ALICE I. ROBOLD, Ball State University, Muncie, Indiana

A study was conducted, during the school year 1962–1963, of the academic backgrounds of 889 college instructors of mathematics content courses for prospective elementary-school teachers. The instructors taught in the fifty states, the District of Columbia, and Puerto Rico.

Some significant findings are as follows:

1. When the instructors were classified according to the interest they indicated in elementary-school teacher education (*least*, *some*, or *greatest*) it was found that instructors who indicated *greatest* interest were less well prepared mathematically, but had better backgrounds in elementary-school teaching, than the instructors who indicated *some* or *least* interest.

2. The percentage of instructors who had been certified to teach or who had taught in elementary school was greater for the group having *least* interest than for the group having *some* interest in elementary-school teacher preparation.

3. Instructors were classified among the seven types of institutions (so classified by the U. S. Office of Education) in which they taught: Liberal Arts and General; Primarily Teacher Preparatory; Liberal Arts and General, and Teacher Preparatory; Liberal Arts and General, Teacher Preparatory, and Terminal Occupational; Professional or Technical and Teacher Preparatory; Liberal Arts and General with One or Two Professional Schools; and Liberal Arts and General with Three or More Professional Schools. The Primarily Teacher Preparatory institutions had a higher percentage of instructors in the group with *least* interest in elementary-school teacher education than any other type of institution.

4. Among the seven types of institutions in which the instructors were employed, Primarily Teacher Preparatory institutions had the lowest percentage of instructors with elementary-school teaching experience; while Liberal Arts and General, and Teacher Preparatory institutions had the highest percentage of instructors with elementary-school teaching experience.

5. Instructors who had been certified to teach in elementary school and instructors who had taught in elementary school had significantly lower means than the other instructors for the number of college mathematics semester credit hours earned (both undergraduate and graduate) and for the number of college mathematics courses (both undergraduate and graduate) they felt competent to teach.

6. Among institutional size groups, instructors in institutions with enrollments from 15,000 to 19,999 had highest means for college mathematics semester credit hours earned and the number of college mathematics courses they felt competent to teach.

7. Fifteen per cent of the instructors had earned less than 40 college mathematics semester credit hours.

8. The Masters degree in mathematics or its equivalent (30 graduate mathematics semester credit hours) had been earned by 522, or 58.7 per cent of the instructors. In this study, 438, or 49.3 per cent of the instructors had earned a Masters degree (or its equivalent) in mathematics and indicated *some* or *greatest* interest in the preparation of elementary-school teachers. Sixty-four, or 7.2 per cent of the instructors had earned a Masters degree (or its equivalent) in mathematics and indicated *greatest* interest in elementary-school teacher education.

9. Instructors who had earned at least 40 college mathematics semester credit hours and indicated *some* or *greatest* interest in elementary-school teacher education comprised 593, or 66.7 per cent of the instructors in the study.

10. Thirty-eight, or 4.3 per cent of the instructors had all of the following characteristics:

Some or *greatest* interest in elementary-school teacher education.
Thirty or more graduate mathematics semester credit hours.
Certification for elementary-school teaching.
Elementary-school teaching experience.

Of this group, only nine (1.0 per cent of the instructors in the study) had *greatest* interest in the preparation of elementary-school teachers.

These findings seem to imply a lack of appropriate preparation of college instructors of mathematics for prospective elementary-school teachers.

There is apparently a need to develop academic programs to prepare instructors who have not only an adequate background in mathematics, but also an understanding of the elementary school and an interest in preparing its teachers.

IMPROVING INSTRUCTION IN GRADUATE AND UPPER-LEVEL UNDERGRADUATE MATHEMATICS

RICHARD KUECHLE, Foothill College, Los Altos Hills, California

In recent years the teaching of mathematics in the elementary and secondary schools has been examined very closely, as has the training of the teachers at those levels. Many changes have been suggested, some of which have been implemented, with more changes to come in the future. While attention has been focused on these areas, less has been given to improving the quality of instruction in mathematics at the university level. From September 1962 until August 1964 I was a graduate student of mathematics at two large state universities, and thus in a favorable position to report on the teaching ability of mathematics professors. My two years of study were in National Science Foundation Academic Year Institutes, culminating in a master's degree in mathematics at the second of the two universities.

First I attended University X for two semesters and the summer session, and at the end of my year there I was disturbed about the quality of teaching,

but thought that perhaps it was a condition peculiar to this university. The federal grant designated University Y for my second year of study, thus giving me the chance to compare the quality of instruction at the two schools. After studying at University Y for one year I reached the same conclusion I had at University X, only this time I was more deeply disturbed, because now I began to wonder about the quality of teaching in the mathematics departments of all universities.

Before discussing any of my observations at the two univerisities, I will set forth, as a basis for comparison, my concept of the pedagogical role of a mathematics professor who teaches graduate and upper-level undergraduate courses. He assumes the following responsibilities:

(1) To prepare a well-organized body of mathematical information that will be presented to the class over the semester.

(2) To present this body of information in an efficient manner designed to assist the students in learning the material.

(3) To prepare a list of homework exercises which thoroughly cover all of the emphasized points in the course, and to assign these at frequent, regular intervals. The exercises which are the most difficult for the students and those which are the most important to the development of the course should be discussed in class.

(4) To plan an evaluation program and to prepare tests which are valid and reliable in measuring a student's achievement in the course.

(5) To create a class climate in which students are stimulated to follow the presentation closely and are encouraged to ask questions.

Point 5 above is more important than one might think. Findings from the psychological theories of learning (conditioning, connectionism, and field psychology) all indicate that the learner must be mentally active and that his thoughts must lead to a successful solution if learning is to take place [1]. This raises the question of how much learning takes place in a class where the students take notes as fast as they can while the professor writes continuously at the board with his back to the audience. This method of teaching might have been defensible before the invention of duplicating machines, but not now.

With the above five points as a basis for comparison my observations of the 14 professors at Universities X and Y can now be presented. All of the 14 held the Ph.D. degree in mathematics, and in rank they ranged from assistant professor to full professor, with one of them being the head of the department. They include names familiar to any serious student of mathematics and mathematics education.

It is interesting to note the frequency with which certain faults occurred. See the following chart in which the poor teaching practices are tabulated.

In judging teaching effectiveness one should not assign equal weight to each of the points listed in this chart, but it is interesting to note that if I were to name the three best teachers in terms of overall performance (and this includes some factors not included in this chart), they would be Professors A, B, and H,

Teaching Faults / Professor	I. Inefficient & Ineffective Use of Class Time						II. Failure to Effectively use Homework as a Learning Device				III. Use of Poor Techniques of Evaluation			
	A	B Poor Presentation of the Day's Lesson				C	A	B	C	D	A	B	C	
		1.	2.	3.	4.									
	Poor Organization	Too vague	Digresses on irrelevant topics	Teaches as if duplicating machines do not exist	Proceeds at too rapid a pace	Conducts a passive class	No core of homework which thoroughly covers the course	No frequent regular assignments	Too little discussion of solutions to homework problems	Poor use of homework to help teach new concepts	Uses too few evaluation instruments	Too much reliance on work done outside of class	Use of Invalid Tests	Totals
A										√	√	√		3
B										√				1
C	√	√				√				√				4
D	√	√			√	√	√	√	√	√				8
E		√		√	√	√	√	√	√	√				8
F	√					√	√			√	√			5
G	√	√	√				√	√		√	√	√		8
H					√	√				√				3
I						√	√	√	√	√		√	√	7
J		√				√	√		√	√			√	6
K				√				√	√	√				4
L		√		√			√			√				4
M	√	√	√		√	√				√	√	√		8
N	√	√		√	√	√		√	√	√	√	√		10

in that order. Each of these men has three or fewer marks against him. The two worst teachers, Professors M and D, have eight marks each, while Professor N, having 10 marks, was only a slightly better teacher. There are seven professors whom I would classify as poor teachers, and they are C, D, E, F, G, M, and N. Of the teaching faults listed on the chart there is one which shows a high degree of correlation with these seven teachers, and that is I-A, "Poor Organization of the Course and Poor Organization of the Daily Presentation." Of these seven, only Professor E does not have this mark against him, and *none* of the other professors (who could be divided into groups of *good* and *fair* teachers) have this mark against them. I do not believe that a subject which has such a high degree of organization within itself can be effectively taught in a disorganized manner. Other suggestions for improving instruction are implied by the list of faults in the chart, but beyond that I have some additional comments.

The first of these has its basis in the premise that a teacher must make the best use of the class time at his disposal. Having accepted this, then I believe that the professor must do more in his class than merely present information. If information-giving were the only reason for holding classes, then for the most part classes could be eliminated, for the students could get the information from the text, or if there were no text for the course, the professor could have sets of notes duplicated for distribution to the class, and the students could study them

and report for examinations at specified times. To use class time effectively, the professor must give the students *guidance* in learning the new material. More than the bare skeleton of information must be presented to the class, and it is at this point where many professors fail as teachers. They see their role in the classroom to be that of lecturers (information-giving) only.

A good teacher should, in addition to giving information, emphasize the most important points in a presentation, uncover relationships that the student might overlook otherwise, and check to see that the class is comprehending the material. If classes are worth holding at all, then learning *must* take place in them, and this implies that information-giving and guidance must both take place in the classroom.

Any student who has the ability to earn a master's degree or higher in mathematics certainly has the ability to sit down and study the text until eventually he understands it. Why, then, am I making this fuss about guiding the student in class? There are two reasons. The first is that if the professor furnishes guidance in class, the student will be able to master the topic in less time than if he were left to his own devices. The graduate student has a heavy enough load without being delegated the added responsibility of teaching himself topics that the professor failed to present effectively in class. The second reason is one of motivation. A student who receives no guidance from his professor will probably do more floundering about before mastering the lesson than will a student who is effectively guided. I think that the former student will eventually become negatively motivated after enough repetitions of this experience and will be less likely to succeed.

The problem of improving instruction in university mathematics is a complex one and hence, there can be no one simple solution. I think that the first step in any solution will be to convince the professors and administrators that there is a need to improve the quality of instruction in mathematics, because if they do not see this, then they will not approve any plan to remedy the situation. Hence, the fact that the problem exists and needs a solution must be brought to the attention of those who have the power to implement the necessary changes. University administrators exert pressure on faculty members to engage in research, and one of the motives behind this, besides the obvious one of helping to push forward the frontier of knowledge, is to bring recognition to the university. Is it not equally important that the university be recognized as an institution where a student can get a fine education by excellent teachers? A research-oriented university must guard against losing touch with the individual student. There is a growing awareness of this problem, and one graduate school dean who has recognized it is Dean Richard Armitage of Ohio State University [2]:

"Please notice that I have not, in my characterization of the graduate professor, said that his strong research motivation is improper. I am merely saying that he is at least partly at fault . . . for the strong distrust on the departmental level of anyone who devotes a great deal of time to undergraduate teaching, to counseling students, to writing of textbooks, to participation in honors programs and to the design of exciting new courses in general education."

Some suggestions which have been offered by others for improving the conditions noted in this paper include the following: rewarding good teachers with increases in pay and promotions in rank; and asking that those Ph.D. candidates who plan to become professors take courses in the teaching of mathematics and do some student teaching at the university level.

Another possibility is to set up in-service courses or seminars in which teaching methods would be discussed. These could be intra-departmental, or they could include special consultants from the psychology or education departments. The promise of salary increases afterwards might increase the number of participants in such a program.

Research is needed into the problems which are peculiar to the teaching of university mathematics. Some universities have established centers whose purpose is to study ways of improving the teaching of all disciplines. The advent of closed circuit television, programmed texts, and other new teaching devices may hold some of the answers to the problems facing university mathematics professors.

References

1. Howard F. Fehr, Theories of learning related to the field of mathematics, *The Learning of Mathematics. Its Theory and Practice*, 21-st Yearbook of the National Council of Teachers of Mathematics, Washington, D. C., 1953, p. 31.

2. Richard Armitage, The quality of our undergraduate teaching is suffering, *The Ohio State University Monthly*, April, 1964, p. 13.

A BRIEF DICTIONARY OF PHRASES USED IN MATHEMATICAL WRITING

H. PÉTARD, Society for Useless Research

Since authors seldom, if ever, say what they mean; the following glossary is offered to neophytes in mathematical research to help them understand the language that surrounds the formulas. Since mathematical writing, like mathematics, involves many undefined concepts, it seems best to illustrate the usage by interpretation of examples rather than to attempt definition.

ANALOGUE. This is an a. of: I have to have *some* excuse for publishing it.

APPLICATIONS. This is of interest in a.: I have to have *some* excuse for publishing it.

COMPLETE. The proof is now c.: I can't finish it.

DETAILS. I cannot follow the d. of X's proof: It's wrong. We omit the d.: I can't do it.

DIFFICULT. This problem is d.: I don't know the answer. (Cf. Trivial.)

GENERALITY. Without loss of g.: I have done an easy special case.

IDEAS. To fix the i.: To consider the only case I can do.

INGENIOUS. X's proof is i.: I understand it.

INTEREST. It may be of i.: I have to have *some* excuse for publishing it.

INTERESTING. X's paper is i.: I don't understand it.

KNOWN. This is a k. result but I reproduce the proof for the convenience of the reader: My paper isn't long enough.

Langage. Par abus de l.: In the terminology used by other authors. (Cf. Notation.)

NATURAL. It is n. to begin with the following considerations: We have to start somewhere.

NEW. This was proved by X but the following n. proof may present points of interest: I can't understand X.

NOTATION. To simplify the n.: It is too much trouble to change now.

OBSERVED. It will be o. that: I hope you have not noticed that.

OBVIOUS. It is o.: I can't prove it.

READER. The details may be left to the r.: I can't do it.

REFEREE. I wish to thank the r. for his suggestions: I loused it up.

STRAIGHTFORWARD. By a s. computation: I lost my notes.

TRIVIAL. This problem is t.: I know the answer. (Cf. Difficult.)

WELL-KNOWN. This result is w.: I can't find the reference.

Exercises for the student: Interpret the following.

1. I am indebted to Professor X for stimulating discussions.
2. However, as we have seen.
3. In general.
4. It is easily shown.
5. To be continued.

This article was prepared with the opposition of the National Silence Foundation.

ADMINISTRATIVE GUIDE FOR SCHOOL PROGRAMS

The National Council of Teachers of Mathematics, in cooperation with other national professional organizations, has recently published a brochure entitled, *Administrative Responsibilities for Improving Mathematics Programs.* Representatives of the American Association of School Administrators, Association of Curriculum Development, and the National Association of Secondary School Principals cooperated in preparation of the publication. The purposes of the publication were to identify the kinds of organization that make possible sound administration of good mathematics programs, to outline the principles of curriculum planning and revision, and to define responsibilities of various school personnel in the revision process. A wide distribution is guaranteed through the American Association of School Administrators and the National Council of Teachers of Mathematics. Both groups are distributing the bulletin without cost. Members representing the National Council of Teachers of Mathematics included: Mildred Keiffer, mathematics supervisor, Cincinnati; Frank B. Allen, Lyons Township High School and Junior College; and Charles R. Hucka, Eugene, Oregon Public Schools.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE, Bloomfield College

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, Bloomfield College, Bloomfield, N. J. 07003. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

All solutions of Elementary Problems should be sent to A. E. Livingston, University of Alberta, Edmonton, Canada. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before June 30, 1966.

E 1855. *Proposed by L. M. Young, San Fernando Valley State College*

Let λ be a given (real or complex) number. Solve the system:

$$x_{n+1} = x_n^2 - \lambda \qquad (n = 1, 2, 3, 4),$$

where $x_5 = x_1$, subject to the condition $x_3 \neq x_1$.

E 1856. *Proposed by S. W. Williams, Lehigh University*

Given n any positive integer, prove the identity

$$\sum_{k=1}^{n} [kx] = n[x] + \sum_{i=1}^{n-1} \sum_{k=1}^{i} \left[x + \frac{k}{i+1} \right],$$

where, as usual, $[x]$ is the greatest integer in x.

E 1857. *Proposed by Steven Rosencrans and James Potter, Massachusetts Institute of Technology*

Show that the solutions λ of $\tan \lambda = a\lambda/(\lambda^2 + c)$, where a is real and c is positive, are either real or (pure) imaginary.

E 1858. *Proposed by S. E. Dickson, University of Nebraska*

Assume the usual properties of the real numbers except for commutativity of addition. Given that the equation $x+y=y+x$ holds for real numbers x and y less than 10^{-26} apart, show that this latter restriction can be removed.

E 1859. *Proposed by Judith H. Hallman, University of North Carolina and Western Electric Co.*

Let $k_1, k_2, \cdots, k_n$ be any real numbers different from zero, and let $A = (a_{ij})$

be the matrix $a_{ij}=k_i/k_j$. Find the n characteristic roots of A.

E 1860. *Proposed by Edward Thorp and Fred Richman, New Mexico State University*

Consider the sequences n^n, $n^{(n^n)}$, etc., and $n!$, $n!!$, $n!!!$, etc., for a fixed integer $n \geqq 3$. Do the terms of one sequence eventually become larger, and stay larger, than the corresponding terms of the other? Suppose, instead, that we compare the corresponding terms of n^n, $n^{(n^n)}$, etc., and $n!!$, $n!!!$, etc., and ask the same question.

E 1861. *Proposed by T. R. Curry, Bay Shore, N. Y.*

Given a triangle ABC with the usual notation for sides, angles, area, circumradius and inradius, prove the following inequalities, with equality holding if and only if the triangle is equilateral:

$$4\Delta\sqrt{3} \leqq \frac{9abc}{a+b+c}, \tag{1}$$

$$2\Delta\sqrt{3} \leqq 9rR, \tag{2}$$

$$\cot A + \cot B + \cot C \geqq \frac{\sqrt{3}(a^2+b^2+c^2)(a+b+c)}{9abc}. \tag{3}$$

$$\csc A + \csc B + \csc C \geqq 2\sqrt{3}. \tag{4}$$

E 1862. *Proposed by Michael Fried, University of Michigan*

Let $A=\{a_i\}_{i=1}^{\infty}$ be an ordered set of distinct positive integers. Let

$$P(x) = \prod_{i=1}^{\infty} (1 + x^{a_i}) = \sum_{i=0}^{\infty} N_A(i)x^i.$$

It is well known that $N_A(n)$ is the number of ways n may be expressed as a sum of distinct elements from A.

$N_A(n)$ is said to be unbounded if $\lim_n \sup N_A(n) = \infty$.

$N_A(n)$ is said to be uniformly unbounded if $\lim_n N_A(n) = \infty$.

(a) Find necessary and sufficient conditions on the set A so that $N_A(n) > 0$ for all n.

(b) Find necessary and sufficient conditions on the set A so that $N_A(n)$ is unbounded.

(c) Does there exist a set A for which $N_A(n)$ is unbounded but not uniformly unbounded?

E 1863. *Proposed by R. G. Albert, West Newton, Mass.*

The size and shape of a conic with positive eccentricity are determined by two parameters, the distance d from a focus to the nearest directrix and the eccentricity e. Evaluate the radius of the osculating circle at a vertex as a function of d and e.

E 1864. *Proposed by R. E. Mikhel, Tri-State College, Angola, Ind.*

If G_i is an abelian group of involutoric matrices ($A^2=I$), prove that the order of G_i is of the form 2^n.

SOLUTIONS OF ELEMENTARY PROBLEMS

Expected Distance between the Vertices of a Dodecahedron

E 1752 [1965, 75]. *Proposed by Arthur Engel, Stuttgart, Germany*

A bug is crawling on the edges of a dodecahedron. Each time it comes to a vertex it chooses with equal probability one of the three edges which end in that vertex. What is the expected distance it covers in order to get from a vertex A to any other vertex B? ($A=B$ is not excluded.)

Solution by T. Teichmann, General Dynamics Corporation, San Diego. Starting from any vertex "O" on a dodecahedron, the remaining 19 vertices can be divided into 5 classes, characterized by the minimum number of steps required to reach them from "O." (This fact and the following combinations, are easily seen by examining the associated isomorphic planar graph. See, for example, *Graphs and Their Uses* by O. Ore, Random House, p. 29, Fig. 2.5.2.). Class "1" consists of 3 points, Class "2" of 6, Class "3" of 6, Class "4" of 3 and Class "5" of 1. Let n_i be expected number of steps to reach a point in class "i." One then has the following relations (easily obtained from inspection of the graph.)

$$\begin{aligned} n_1 &= \tfrac{1}{3} + \tfrac{2}{3}(n_2 + 1) \\ n_2 &= \tfrac{1}{3}(n_1 + n_2 + n_3 + 3) \\ n_3 &= \tfrac{1}{3}(n_2 + n_3 + n_4 + 3) \\ n_4 &= \tfrac{1}{3}(2n_3 + n_5 + 3) \\ n_5 &= n_4 + 1. \end{aligned}$$

Solving these equations successively, starting with the last, one finds

$$n_1 = 19 \qquad n_2 = 27 \qquad n_3 = 32 \qquad n_4 = 34 \qquad n_5 = 35.$$

If n_0 is the expected number of steps to return to the origin, then $n_0=n_1+1=20$.

Also solved by M. A. Bershad, T. A. Brown, F. P. Callahan, R. E. Glacken, Michael Green, H. S. Hahn, E. L. Magnuson, D. C. B. Marsh, Donald Quiring, Robin Sibson (England), and the proposer.

A Countable Family of Subsets of the Positive Integers

E 1753 [1965, 75]. *Proposed by J. O. Herzog and L. J. Simonoff, Idaho State University*

Let β be a family of subsets of positive integers such that if A, $B\in\beta$, then $A\cap B$ has at most n elements. Prove that β is countable.

Solution by P. L. Manley and M. G. Murdeshwar, University of Alberta, Edmonton, Alberta, Canada. The proposers no doubt meant to require that *distinct* sets $A, B \in \beta$ have at most n elements in common since otherwise each set in β has at most n elements, in which case the conclusion is well known.

With the modified hypothesis, let E be the set of all positive integral $(n+1)$-tuples, and define $f: \beta \to E$ as follows: If $A \in \beta$ consists of $a_1 < a_2 < a_3 < \cdots$, then

$$f(A) = (a_1', a_2', \cdots, a_{n+1}')$$

with $a_k' = a_k$ if $a_k \in A$ and $a_k' = 0$ if A has less than k members. Clearly f is one-to-one, and E is countable.

If "n" is replaced by "finite," β is not necessarily countable. For instance, for each real number $x \in (0, 1)$ define a set F_x consisting of $2^n(2[nx]+1)$, $n = 1, 2, 3, \cdots$. The square brackets denote the usual largest-integer function. Then the family $\beta = \{F_x | x \in (0, 1)\}$ is uncountable.

Also solved by Robert Bernstein, W. H. Bonney, J. L. Brown, Jr., E. O. Buchman, Jim Campbell, J. E. Connett, Oliver Costlich, R. W. Demming, R. B. Eggleton (Australia), J. D. Featherstone, N. J. Fine, R. W. Gilmer, Jr., H. S. Hahn, D. M. Hancasky, Denis Hanson and David Klarner (jointly), A. G. Heinicke, G. A. Heuer, E. S. Langford, C. C. Lindner, D. C. B. Marsh, Maurice Navel and J. J. Zeltmacher (jointly), James Nussbaum, Martin Peres, Donald Quiring, J. G. Rau, Stewart Robinson, L. E. Rudinski, P. J. Saavedra, M. S. R. K. Sastry, Barry Simon, Richard Sinkhorn, G. P. Speck, R. L. Syverson, J. P. Thomas, Guy Torchinelli, B. R. Toskey, R. C. Vile, Jr., Mrs. M. R. Wiscamb, and the proposers.

The proposers also observe that the conclusion does not obtain when "n" is replaced by "finite" (W. Sierpinski, *Cardinal and Ordinal Numbers*, Pansat wowe Wydawnictevo Nankowe, Warsaw (1958) p. 77)..

Several of the solvers solved only the original version of the problem.

A Well-known Continuant

E 1754[1965, 75]. *Proposed by V. R. Rao Uppuluri, Oak Ridge National Laboratory*

Let $p, q > 0$ and $p+q=1$, and let a_m denote the value of the mth order determinant

$$\begin{vmatrix} 1 & -q & 0 & 0 & 0 \cdots & 0 & 0 & 0 \\ -p & 1 & -q & 0 & 0 \cdots & 0 & 0 & 0 \\ 0 & -p & 1 & -q & 0 \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot\cdot\cdot\cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 \cdots & -p & 1 & -q \\ 0 & 0 & 0 & 0 & 0 \cdots & 0 & -p & 1 \end{vmatrix}$$

with $a_0 = 1$ $(=a_1)$. Show that $a_m = (p^{m+1} - q^{m+1})/(p-q)$.

Solution by Douglas Lind, University of Virginia. The conclusion of the problem should of course be $a_m = \sum_{k=0}^{m} p^{m-k}q^k$ to take care of the case where $p=q$,

and our argument shows that the requirement of positivity of p and q is not needed.

It is known that the nth-order continuant $A_n = \det(a_{ij})$ with $a_{ii}=c$, $a_{i,i+1}=a$, $a_{i-1,i}=b$, and $a_{ij}=0$ otherwise, has the value $(n+1)c^n/2^{n+1}$ when $c^2=4ab$ and

$$\frac{(c+\sqrt{c^2-4ab})^{n+1}-(c-\sqrt{c^2-4ab})^{n+1}}{2^{n+1}\sqrt{c^2-4ab}}$$

if $c^2 \neq 4ab$. (See the solution to Problem B-13, Fibonacci Quarterly, 1 (1963) p. 79, or expand A_n by minors according to the elements of the last row or column to get

$$A_n = cA_{n-1} - abA_{n-2}$$

for $n=2, 3, 4, \cdots$ and then apply the usual techniques for solving second-order linear difference equations.). The required result follows upon taking $c=1$, $b=-p$, and $a=p-1=-q$.

Also solved by J. C. Abad, A. N. Aheart, J. C. Ahuja, D. P. Ambrose (Basutoland), J. T. Anderson, J. H. Avila, W. H. Bailey, W. T. Bailey, Raymond Balbes, C. W. Barnes, Robert Bart, Homer Bechtell, John Beidler, P. M. Berry, M. A. Bershad, Marjorie Bicknell, D. A. Blaeuer, W. R. Boland, W. G. Brady, J. L. Brenner, C. A. Bridger, J. L. Brown, Jr., Bruce Bursack, J. A. Burslem, F. A. Butter, Jr., Jim Campbell, R. E. Chandler, M. R. Chowdhury (Germany), Allan Chuck and Peter Goldstein (jointly), Glenn Clark, C. G. Clarridge, Edward Collett, D. I. A. Cohen, R. J. Cormier, Jim Dillinger, R. L. Disney, Ragnar Dybvik (Norway), E. S. Eby, R. B. Eggleton (Australia), J. W. Ellis, W. O. Egerland, N. Ersec, N. J. Fine, Gary Ford, Tom Foregger, E. Frank, Philip Fung, P. K. Garlick, Robert Gouw, Frederick Gray and R. L. Price (jointly), S. H. Greene, Louise S. Grinstein, Cornelius Groenewoud, Harry Guess, H. S. Hahn, D. M. Hancasky, Eldon Hansen, R. J. Harshbarger, J. Z. Hearon, A. G. Heinicke, J. C. Hickman, Francis Higman, Stephen Hoffman, R. G. Hoffmann, K. L. Jensen, Rick Jones, Terry Kempf, M. S. Klamkin, W. J. Klimczak, Dale Klungness, Kenneth Kramer, C. D. La Budde, E. S. Langford, Rene Laumen (Belgium), T. J. Lee, Horst-Eckert Lehmann (Germany), Steve Ligh, C F. McLaren, Robert Maas, E. L. Magnuson, Andrzej Makowski (Poland), D. C. B. Marsh, Helen Merkel, Michael Minkoff, Cleve Moler, J. E. Mueller, M. G. Murdeshwar, G. L. Musser, W. M. Nair, James Newton, Dave Nixon, James Nussbaum, David Palmer, F. D. Parker, Robert Patenaude, C. B. A. Peck, Walter Penney, J. M. Perry, J. R. Porter, Kenneth Quandt, D. B. Ramsay, J. M. Recht, Simeon Reich, Judith Richman, V. K. Rohatgi, Bernard Rosner, W. M. Sanders, M. S. R. K. Sastry, S. W. Saunders, P. A. Scheinok, E. M. Scheuer, Gerald Schrag, A. Sharma, Robin Sibson (England), R. Sivaramakrishnan (India), Barry Simon, Beverly Smith, Mitchell Snyder, Sidney Spital, G. S. Sprouse, M. N. S. Swamy, R. L. Syverson, R. P. Tapscott, J. J. Tattersall, B. R. Toskey, Elias Toubassi, A. M. Vaidya, C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), Emanuel Vegh, Julius Vogel, Gerald Werner, C. R. Williams, P. C. Yang, K. L. Yocom, Charles Ziegenfus, Alexander Zujus, and the proposer.

Most of these solutions were based upon the recurrence relation

$$a_n = a_{n-1} - pqa_{n-2} \ (n=2, 3, 4, \cdots)$$

obtained as in Lind's solution above followed by an application of the Second Principle of Mathematical Induction. A goodly number of the solvers pointed out that the positivity of p and q could be relaxed as in the solution given here, or to $p \neq q$.

Bridger calls attention to the classical gambler's ruin problem: W. Feller, *Probability Theory and Its Applications*, vol. 1, Wiley (1952) p. 282.

Several references to the evaluation of determinants of the form in this problem and to similar problems were given: D. K. Faddeev and I. S. Sominskii, *Problems in Higher Algebra* (translated by J. L. Brenner), Freeman (1965), problem 217; Fibonacci Quarterly, 1 (1963), problem B-46; F. B. Hildebrand, *Methods of Applied Mathematics*, Prentice-Hall (1952) p. 365, problem 54; Thomas Muir, *A Treatise on the Theory of Determinants* (revised by W. H. Metzler), Dover (1960), Chapter XIII, and *History of Determinants*, Vol. 2, Dover, p. 429; O. Perron, *Die Lehre von den Kettenbruchen*, Teubner (1954), pp. 7–9; I. V. Proskuriakov, *Collection of Problems of Linear Algebra* (Russian), Moscow (1957), problem 366.

Squarefree Perfect Numbers

E 1755 [1965, 75]. *Proposed by A. M. Vaidya, Pennsylvania State University*

Prove that 6 is the only squarefree perfect number.

I. *Solution by F. A. Butter, Jr., Los Angeles, California.* If N is squarefree and perfect, then $N=p_1p_2\cdots p_n$ where each p_k is a prime, $n>1$ (no prime is perfect), $1<p_1<p_2<\cdots<p_n$,

$$\sigma(N)=\prod_{i=1}^{n}\frac{p_i^2-1}{p_i-1}=(p_1+1)(p_2+1)\cdots(p_n+1),$$

and $\sigma(N)=2N$. For odd squarefree N, $4\mid\sigma(N)$ but $4\nmid N$, so N (and, hence, p_1) must be even. But then $2^n\mid\sigma(N)$ and $4\|2N$, so that $n=2$ and $4p_2=3(p_2+1)$; i.e., $p_2=3$ and $N=6$ (which is of course a perfect number).

II. *Solution by Robert W. Prielipp, University of Wisconsin.* Suppose that N is a perfect number. If N is even, then $D=2^{n-1}(2^n-1)$ for some positive integer n with 2^n-1 a prime (L. E. Dickson, *History of the Theory of Numbers*, vol. 1, Chelsea, (1952) p. 19). But then $n\geqq 2$. For N to be squarefree, $n=2$ and $N=6$. Now, an odd perfect number (if there is one) is of the form $p^{4t+1}k^2$ with $p=4n+1$ a prime (*loc. cit.*), and this is never squarefree, so there is no odd squarefree perfect number.

Also solved by J. C. Abad, Norma E. Abel, A. N. Aheart, R. G. Albert, J. T. Anderson Joseph Arkin, C. W. Barnes, Jack Barone, Homer Bechtell, Sister Marion Beiter, Robert Bernstein, M. A. Bershad, W. E. Bodden, W. R. Boland, W. H. Bonney, Dale Burnett, J. A. Burslem, Mannis Charosh, John Christopher, Allan Chuck and Peter Goldstein (jointly), D. I. A. Cohen, F. B. Crippen, D. M. Crystal, W. G. Dotson, Jr., R. B. Eggleton (Australia), J. W. Ellis, N. J. Fine, David Finkel, R. L. Forbes and A. S. B. Holland (jointly), P. K. Garlick, R. W. Gilmer, Jr., A. A. Gioia, Michael Goldberg, Jerry Goodman, H. S. Hahn, D. M. Hancasky, M. H. Hayamizu, J. A. H. Hunter, M. S. Klamkin, C. D. La Budde, E. S. Langford, W. G. Leavitt, Douglas Lind, C. C. Lindner, Charles McCracken, Robert Maas, Andrzej Makowski (Poland), D. C. B. Marsh, Mr. and Mrs. D. L. Muench (jointly), J. B. Muskat, Robert Patenaude, C. B. A. Peck, Harsh Pittie, Edith Rasmussen, Simeon Reich (Israel), D. P. Roselle, Fred Rosenblum, M. S. R. K. Sastry, P. A. Scheinok, J. J. Tattersall, G. C. Thompson, Guy Torchinelli, C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), Emanuel Vegh, Julius Vogel, John Waddington, L. J. Warren, Barbara A. Welsh, J. E. Wilkins, Jr., K. L. Yocom, Charles Ziegenfus, Alexander Zujus, and the proposer. Partial solutions by J. W. Baldwin, Marjorie Bicknell, R. J. Cormier, R. E. Harper, R. I. Matthews, and Sam Newman.

Colored Maps on Spheres

E 1756 [1965, 76]. *Proposed by J. P. Ballantine, University of Washington*

Maps consist of lines, vertices, and regions on a sphere. Let there be precisely 3 lines at each vertex. A region is called odd (even) if the number of lines in its boundary is odd (even). If the regions of a map are each assigned one of the colors 1, 2, 3, 4 with no two adjacent regions of the same color, then the map is said to be colored. Prove: If a map is colored, the number of odd regions colored any two colors, say 3 and 4, is even. (Cf. solution E1667 [1965, 81].)

I. *Solution by D. I. A. Cohen, Student, Princeton University.* Let the number of odd regions colored 3 or 4 be N. Recolor each region of color 4 with color 3, and remove the boundary lines between any two regions of the same color. By considering in turn non-adjacent and adjacent odd regions and adjacent odd and even regions (originally colored 3 or 4), it is clear that the number N' of odd regions now colored 3 has the same parity as N. But now each region colored 3 is bounded by countries colored alternately 1 and 2 since we have 3 edges at each vertex of such a region except at the two vertices left upon deletion of a side originally between an odd and an even country colored 3 and 4. But this is impossible for an odd region, so $N'=0$ and N is even.

II. *Solution by Agnis Kaugars, Kalamazoo College.* A necessary condition for the theorem to hold is that every line end in two (distinct) vertices.

Given a colored map, construct at each vertex a "small" triangle, whose color is uniquely determined. Each triangle constructed changes the parity of the number of odd regions of each color, so the conclusion holds for the resulting map if and only if it held for the original map. The process doubles the number of boundary lines of every region in the original map, so the only odd regions remaining are the triangles. Thus the theorem can be equivalently stated: *If in a colored map we associate with each vertex the one color which does not appear in the three regions meeting there, then the number of vertices associated with any two colors is even.*

Let A and B be any two colors. Call lines which are between regions colored A and B *dividers*. At the ends of each divider are two vertices, neither of which can be colored A or B. By the properties of a colored map, no vertex can be on two dividers. If a vertex is on no divider, then A or B does not appear in the three regions meeting there, and that vertex must be associated with either A or B. Hence there are an even number of vertices associated with neither A nor B, which must thus be associated with one of the other two colors. Since A and B were arbitrarily chosen, the number of vertices associated with any two colors is even, proving the theorem.

Also solved by the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before August 31, 1966.

5360. *Proposed by R. B. Eggleton, Avondale College, Australia*

If μ is the Möbius function, evaluate:

$$S(a) = \sum_{d|a} \mu(d)\mu(a/d) \quad \text{and} \quad S^*(a) = \sum_{d|a} \sum_{m|d} \mu(m)\mu(d/m).$$

5361. *Proposed by W. E. Sewell, Duke University*

Let $P_n(z)$ be a polynomial of degree n in $z=re^{i\theta}$ with only even powers of z, and let $|P_n(z)| \leqq M$ for z on the curve C whose polar equation is $r=|\cos \frac{1}{2}\theta|$. Show that, for z on C, $|P_n'(z)| \leqq 3Mn$.

5362. *Proposed by John de Pillis, San Francisco State College*

Let $\mathfrak{X}$ be the n-dimensional vector space of n-tuples with complex entries. Suppose $n=p\cdot q$ for some integers p and q. The elements X of $\mathfrak{X}$ can then be represented by the n-tuple

$$X = (x_{11}, x_{12}, x_{13}, \cdots, x_{1q}, x_{21}, x_{22}, \cdots, x_{2q}, \cdots, x_{p1}, \cdots, x_{pq}).$$

The ordering is lexicographic by subscripts. If every $X\in\mathfrak{X}$ is identified with its $p\times q$ matrix (the ijth entry $=x_{ij}$) then for any orthonormal basis $\{X^1, X^2, \cdots, X^n\}$ of $\mathfrak{X}$, show that $\sum_{i=1}^n X^iX^{i*}=q\cdot I_p$, where I_p is the $p\times p$ identity matrix and X^{i*} is the adjoint (conjugate transpose) of X^i.

5363. *Proposed by Joseph Kitchen, Duke University*

Let S be an isometry of a Hilbert space H into itself. Prove that there exists a closed subspace M and a unitary operator U on H such that (a) M^o (the orthogonal complement of M) lies in the range of S, (b) U acts like S on M, and (c) U acts like S^{-1} on M^o.

5364. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

Let $f(x)$ be an integral valued function defined for all positive integers with $0\leqq f(x)\leqq x^a$, $0<a<1$. (a) Prove that the graph of $f(x)$ contains at least k collinear points for any positive integer k. (b) Is the same conclusion true when $0\leqq f(x)\leqq x$?

5365. *Proposed by A. Himmelfarb, Fordham University*

Let R be a ring with identity and n a natural number greater than 1. Show that there exist two nonsingular matrices whose sum is the identity matrix.

5366. *Proposed by H. D. Keesing, University of Wisconsin*

Find an explicit function f on $[0, 1]$ to the real numbers, such that f' is defined on $[0, 1]$ and discontinuous on the rationals in $[0, 1]$.

5367. *Proposed by D. J. Newman, Yeshiva University*

Let $d\mu(x)$ be any finite positive measure and set $\phi(t) = \int_1^\infty \cos xt \, d\mu(x)$. Prove that $\phi(t)$ must have a zero in $[0, \pi]$. Prove also that π cannot be replaced by any smaller number.

5368. *Proposed by Joseph Hammer, University of Sydney, Australia*

A curve of constant width greater than $\pi/(\pi-\sqrt{3})$ contains a lattice point.

5369. *Proposed by Albert Wilansky, Lehigh University*

Let M be the set of infinite matrices of real numbers each of which is row finite and column finite (i.e. each row and column has finitely many nonzero entries).

Let f: $M \to$Reals satisfy $f(I) = 1$, $f(AB) = f(BA)$.

Show that $\{A : f(A) = 1\} \not\subset G$, where G is the set of invertible members of M. (The result is false for finite matrices of fixed order.)

SOLUTIONS OF ADVANCED PROBLEMS

Diagonals of a Polygon

4444 [1951, 422; 1962, 173]. *Proposed by J. H. Braun, Illinois Institute of Technology*

Prove that no three diagonals of a regular polygon of odd order are concurrent at any point other than the vertices.

Comment by Andrzej Makowski, University of Warsaw, Poland. The complete proof has been given by Hermann Heineken, *Regelmässige Vielecke und ihre Diagonalen*, Enseignement Math. (2), 8 (1962) 275–278.

Maximal Nonsingular Subspaces

5027 [1962, 438; 1963, 580; 1963, 1016; 1964, 441]. *Proposed by A. J. Goldman, National Bureau of Standards*

Let $M_n(F)$ be the set of $n\times n$ matrices over the field F, considered as an n^2-dimensional vector space over F. Call a vector subspace of $M_n(F)$ nonsingular if all its nonzero members are nonsingular matrices. Find maximal nonsingular subspaces of $M_n(F)$.

Partial Solution by D. M. Topping and P. A. Fillmore, University of Chicago. The maximum dimension of nonsingular subspaces of the $n\times n$ matrices over a field F is n. In fact, the linear space spanned by any $n+1$ matrices contains a

singular matrix, for the first rows of these $n+1$ matrices are linearly dependent vectors.

This maximum is sometimes attained. If D is an n-dimensional (not necessarily associative) division algebra over F then the regular representation of D yields an n-dimensional nonsingular subspace of the $n\times n$ matrices over F.

A classification of the nonsingular subspaces of the $n\times n$ matrices over F does (as the editor observes) indeed depend on F. John Thompson has pointed out that if there is a nonsingular subspace of dimension n, then there is also a division algebra over F of degree n (this seems to be well-known folklore).

The nontrivial nature of this question becomes apparent when one asks: For what n is there a nonsingular subspace of dimension n in the real $n\times n$ matrices? A celebrated result on real division algebras tells us that n must be 1, 2, 4, or 8. (See J. Milnor, *Some consequences of a theorem of Bott*, Ann. Math., 68 (1958) 444–449 (especially Cor. 1, p. 445).) The problem, however, lacks a complete solution even for the real field. For odd n the maximum dimension is 1, and it is $\geqq 2$, 4, or 8 according as 2, 4, or 8 divides n, (the value n being attained only for $n=1$, 2, 4, 8).

Additional comment by D. M Topping. The maximum dimension of nonsingular subspaces of the $n\times n$ real matrices is one more than the number of independent, singularity-free vector fields on the $n-1$ sphere S^{n-1} (this is well known to algebraic topologists). If $n=(2a+1)2^b$ and $b=c+4d$, where a, b, c and d are integers and $0\leqq c\leqq 3$, then the maximum dimension of nonsingular subspaces of the $n\times n$ real matrices is 2^c+8d. (See J. F. Adams, *Vector fields on spheres*, Ann. Math., 75 (1962) 603–632.)

Generalized Circulants

5220 [1964, 801]. *Proposed by Johann Cigler, the University of Vienna, Austria*

Let G be a finite Abelian group of order n. Let $A=(a_{g,h})$ be an $n\times n$ matrix, indexed with the elements g, h of G in a certain fixed order. The matrix A is called a G-circulant A_μ if there exists a set of complex numbers $\mu=\{\mu(g)\}$, $g\in G$, such that $a_{g,h}=\mu(g^{-1}h)$. It is easy to see that the set $C(G)$ of all G-circulants forms an algebra over the complex numbers.

(1) Consider μ as a measure on G. Show that the mapping $\mu\to A_\mu$ defines an isomorphism between the convolution algebra $M(G)$ of all complex valued measures μ on G and $C(G)$. (2) Determine the eigenvalues and eigenvectors of A_μ. (3) Let $A_\mu\in C(G)$, $\det A_\mu\neq 0$ and let T be an automorphism of G. Define $A_\mu^T=(\mu((Tg)^{-1}h))$. Determine the eigenvalues and eigenvectors of A_μ^T.

Solution by the proposer. (1) The mapping $\mu\to A_\mu$ is one-one and satisfies $\mu+\nu\to A_\mu+A_\nu$ and $c\mu\to cA_\mu$, where c denotes a complex number. Let $\mu*\nu$ be the convolution of μ and ν, i.e. $\mu*\nu(g)=\sum_{l\in G}\mu(l)\nu(l^{-1}g)$. Then $A_{\mu*\nu}=A_\mu A_\nu$ is a direct consequence of the definition of matrix multiplication.

(2) The eigenvectors of A_μ are the characters $\chi(g)$ of G and the corresponding eigenvalues are the numbers $\hat{\mu}(\chi)=\sum_{g\in G}\chi(g)\mu(g)$. This follows from the relation

$$A_\mu \chi \cdot (g) = \sum_{l \in G} \mu(g^{-1}l)\chi(l) = \sum_{l \in G} \mu(l)\chi(gl) = \hat{\mu}(\chi)\chi(g),$$

and the fact that there are exactly n distinct characters. The determinant of A_μ is therefore given by $\det A_\mu = \Pi_\chi \hat{\mu}(\chi)$.

(3) For every character χ of G we have

$$\text{(i)} \qquad A_\mu^T \chi \cdot (g) = \sum_{l \in G} \mu((Tg)^{-1}l)\chi(l) = \sum_{l} \mu(l)\chi(l)\chi(Tg) = \hat{\mu}(\chi)\chi(Tg).$$

Define $U\chi \cdot (g) = \chi(Tg)$. It is clear that U induces a permutation of the set of all characters. Let $k = k(\chi)$ be the order of the cycle to which χ belongs, i.e. the least positive integer k for which $U^k \chi = \chi$. From (i) we conclude

$$\text{(ii)} \qquad (A_\mu^T)^k \chi \cdot (g) = \hat{\mu}(\chi)\hat{\mu}(U\chi) \cdots \hat{\mu}(U^{k-1}\chi)\chi(g).$$

Let $\lambda(\chi)$ be any one of the k values of

$$\{\hat{\mu}(\chi)\hat{\mu}(U\chi) \cdots \hat{\mu}(U^{k-1}\chi)\}^{1/k}.$$

By assumption $\hat{\mu}(\chi) \neq 0$ and therefore also $\lambda(\chi) \neq 0$. Let

$$\nu_\lambda(g) = \sum_{j=0}^{k-1} \lambda^{k-1-j} (A_\mu^T)^j \chi \cdot (g).$$

Then (i) and (ii) immediately imply $A_\mu^T \nu_\lambda \cdot (g) = \lambda \nu_\lambda(g)$, i.e. λ is an eigenvalue of A_μ^T, and $\nu_\lambda(g)$ is the corresponding eigenvector. The nature of the permutation induced by U implies that all of the eigenvalues and eigenvectors may be obtained by this procedure.

Also solved, partially, by D. Ž. Djoković.

Generators for the Addition Group $R \times R$

5261 [1965, 192]. *Proposed by M. Rajagopalan, University of Illinois*

Let R^2 be the complex plane with the usual addition operation. Let S be the set $\{x + i \sin x \mid 0 \leqq x \leqq \pi\}$. What is the additive group generated by S?

I. *Solution by S. J. Sidney, Harvard University.* More generally, let $S = \{x + i \sin x \mid 0 \leqq x \leqq a\}$ for some $a > 0$. Define $f: [0, a/2] \to \mathcal{R}$, the reals, by $f(x) = 2 \sin x - \sin 2x$. Then the image of f contains the interval $[0, b]$ for some $b > 0$. Given $c \in [0, b]$, choose $x \in [0, a/2]$ such that $f(x) = c$. Then $x + i \sin x$ and $2x + i \sin 2x$ are in S and $2(x + i \sin x) - (2x + i \sin 2x) = ic$. It follows that S generates the entire imaginary axis, and consequently the whole plane.

II. *Solution by Necdet Üçoluk, Purdue University.* Let $R(x, y)$ denote the complex plane, so that $S = \{(x, y) \mid x = u,\ y = \sin u,\ 0 \leqq u \leqq \pi\}$. A subset of the additive group generated by S is the set $T = \{(x, y) \mid x = u + v,\ y = \sin u + \sin v,$

$0 \leqq u, v \leqq \pi\}$. So T is the image of f: $R(u, v) \rightarrow R(x, y)$. This function f is continuously differentiable, with Jacobian equal to $\cos v - \cos u$. If $(u_0, v_0) \in R(u, v)$, $u_0 \neq v_0$, then there is an open neighborhood A of (u_0, v_0) such that $f|A$ is one to one; thus $f(A)$ is a nonempty open set contained in T. The subgroup generated by a nonempty open set of the additive group $R(x, y)$ is the entire plane; which solves the problem.

We note (1) in the problem above, S is an arc of the curve $y = \sin x$. The same proof allows us to use instead of S any continuously differentiable path except one which is part of a ray through the origin. (2) It seems reasonable to assert that S may be taken as any continuous path with the exception as in (1). (3) Moreover, if S generates $R(x, y)$ we expect that S contains a continuous path.

Also solved by Robert Bowen, G. W. Day, D. Ž. Djoković (Yugoslavia), G. A. Heuer, R. E. Maas, M. D. Mavinkurve (India), Robin Sibson (England), F. E. Spencer, Jr., W. C. Waterhouse, and the proposer.

The proposer notes that remark (2) in the second solution implies that an arcwise connected additive subgroup of the plane must be $\{0\}$, or a ray, or the entire plane.

Units in Ring Extensions

5262 [1965, 192]. *Proposed by George Bergman, Harvard University*

The Dirichlet Unit Theorem says that, if R is an integrally closed integral domain with finite additive basis, over the integers, then the multiplicative group of units in R is finitely generated, with rank $m-1$ (and torsion subgroup, of course, equal to the group of roots of unity in R); where m is the number of embeddings of R in the complexes, modulo complex conjugacy. (E.g., $m=2$ for $Z[\sqrt{2}]$, $m=1$ for $Z[\sqrt{-2}]$.)

Prove that the condition that R be integrally closed can be omitted without affecting the conclusion.

Solution by the proposer. Suppose that R is of rank n as an additive group. Then $R \cdot Q$ (Q = the field of rationals) is an integral domain containing the rationals and is of dimension n over Q; hence it is a field. Let $\overline{R}$ be the ring of algebraic integers in this field. Then the groups of units of R and $\overline{R}$ are of the same rank; specifically we shall show that some power of any unit $u \in \overline{R}$ is in the group of units of R.

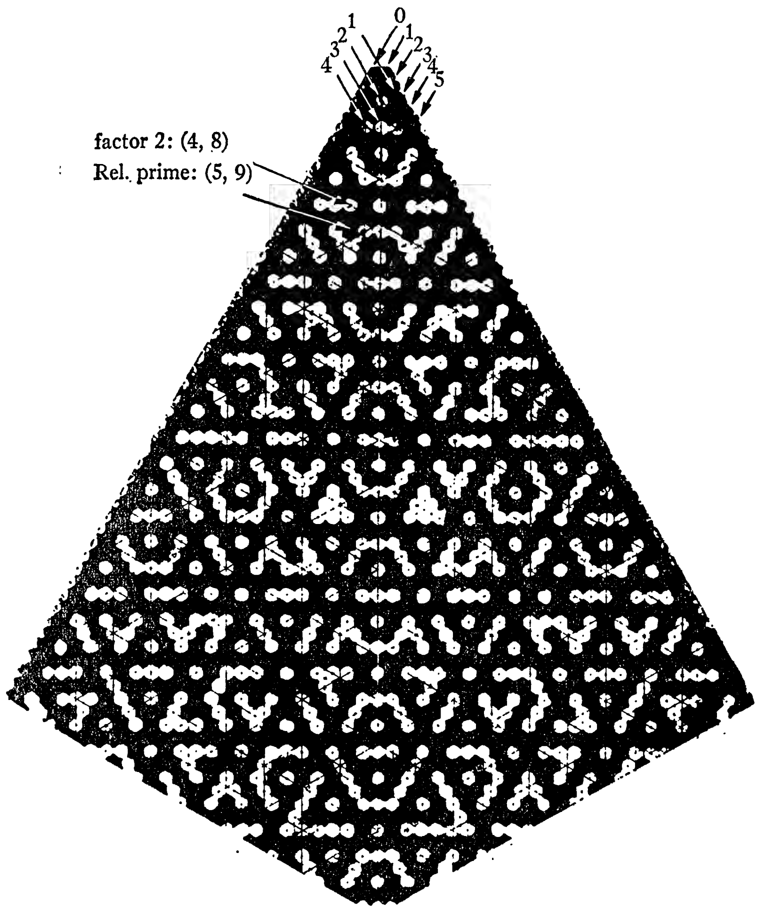
Since $\overline{R}$ and R are both of rank n as additive groups, it follows that for some integer $k > 0$, $k\overline{R} \subset R$. Since $\overline{R}/k\overline{R}$ is finite, the powers of u cannot all be distinct (mod $k\overline{R}$). Suppose $u^p \equiv u^{p+r}$. Multiplying by u^{-p} and u^{-p-r} (which are in $\overline{R}$), we get $u^{\pm r} \equiv 1$ (mod $k\overline{R}$), hence $u^{\pm r} - 1 \in k\overline{R} \subset R$, hence $u^{\pm r} \in R$, and u^r is a unit of R. The result now follows.

Coprime Coordinates in the Regular Tessellation $\{6, 3\}$

5263 [1965, 192]. *Proposed by Art Winfree, Cornell University*

Consider an angular region, of angle 60°, filled with a honeycomb of regular hexagons (that is, one sixth of the complete regular tessellation $\{6, 3\}$). Number

the rows of hexagons 1, 2, 3, $\cdots$ from one side of the angle and again 1, 2, 3, $\cdots$ from the other side, thus assigning a pair of coordinate integers to each hexagon. Prove or disprove that the hexagons with relatively prime coordinates (constituting $6/\pi^2$) of the whole) form a connected set, whereas those whose coordinates have a common factor greater than 1 occur in isolated clusters. (See the figure where the construction is carried out to about 50 by 50.)



Solution by J. P. Altgeld, University of Illinois. The hexagon (34, 21) is an isolated one with relatively prime coordinates. Using the Chinese remainder theorem it is easy to find infinitely many such. Hexagons (x, y) with $x+y$ a prime have relatively prime coordinates and form connected strips that divide the whole into finite pieces.

Thus the hexagons with relatively prime coordinates have one infinite component and infinitely many finite components. The hexagons having a common factor greater than one have infinitely many finite components.

Also solved by T. J. Bruggeman, M. Lieber and W. Weissblum, F. Göbel (Netherlands), and B. M. Stewart.

Lieber and Weissblum report the following hexagons with relatively prime coordinates but all of whose neighbors have coordinates with common factors: (21, 55), (76, 99), (99, 175), (206, 369), (375, 424). Bruggeman asks for necessary and sufficient conditions that for a given hexagon (m, n) we have $(m, n)=d>1$, but for each of the neighbors $(m-1, n)$, $(m, n-1)$, $(m-1, n+1)$, $(m+1, n-1)$, $(m, n+1)$, $(m+1, n)$ we have relatively prime coordinates. The requirement is satisfied if m, n are equal even numbers.

The Triangle Inequality and the Parallelogram Law

5264 [1965, 193]. *Proposed by D. E. Knuth, California Institute of Technology*

Let V be a vector space over the real numbers, and let the operator $\|x\|$ be defined for all $x\in V$, satisfying the following conditions:

(1) $\|x\|$ is a nonnegative real number.

(2) $\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2$ (parallelogram law).

Prove that the triangle inequality holds:

$$\|x+y\| \leqq \|x\| + \|y\|.$$

I. *Solution by E. O. Buchman, University of California, Los Angeles.* Suppose to the contrary that we have a, b such that $\|a+b\|>\|a\|+\|b\|$. Then there is an $\epsilon>0$ such that we also have

$$\|a+b\| > \|a\| + \|b\| + \epsilon. \tag{1}$$

Write $c=a+b$, and suppose, without loss of generality, that $\|a\|\geqq\|b\|$. We shall obtain the absurdity that for every positive integer m, $\|c\|>m\epsilon$. To achieve this, we proceed to prove, with the parallelogram law and the assumptions of this paragraph, that

$$\|a\| > \|b\| + \|a-b\| + \epsilon. \tag{2}$$

Proof of (2). (a) Since $\epsilon>0$, $\|b\|\geqq 0$, $2\|a\|^2+2\|b\|^2=\|a+b\|^2+\|a-b\|^2$, and $[\|a\|+\|b\|+\epsilon]^2<\|a+b\|^2$, it follows that

$$(\|a\| - \|b\| - \epsilon)^2 > \|a-b\|^2.$$

(b) From the additional assumption $\|a\|\geqq\|b\|$ it follows that $\|a\|-\|b\|>\epsilon$. To prove this last assertion, assume to the contrary, that $\|a\|-\|b\|\leqq\epsilon$ and write $\eta=\|a\|-\|b\|$, $0<\eta\leqq\epsilon$. It follows, therefore, that $\|a\|+\|b\|+\eta<\|a+b\|$ and by part (a) just above, we must also have $(\|a\|-\|b\|-\eta)^2>\|a-b\|^2$ or $0>\|a-b\|^2$, which is impossible. (2) now follows.

Combining (1) and (2), we see that $\|c\|(=\|a+b\|)>2\epsilon$. Moreover, any statement of this form which is proved for $\|c\|$ is similarly true for all $\|a\|$, since $\|a\|$ satisfies an inequality equivalent to the inequality for $\|c\|$. Hence $\|a\|>2\epsilon$ which implies $\|c\|>3\epsilon$ which implies $\|a\|>3\epsilon$ and so forth. Thus, for all positive m, $\|c\|$ (and $\|a\|$ too) $>m\epsilon$. From this impossible implication, we conclude that the parallelogram law does indeed imply the triangle inequality.

II. *Solution by W. G. Dotson, Jr., G. C. Marshall Space Flight Center, Huntsville, Alabama.* We observe first that $\|0\|=0$ and $\|-x\|=\|x\|$. Define an inner product (x, y) by

$$(x, y) = \tfrac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

We have immediately $(x, x)=\|x\|^2$, and $(x, y)=(y, x)$. Computing from the parallelogram law (several times) we obtain $(x+z, y)=(x, y)+(z, y)$. It then follows that $(rx, y)=r(x, y)$ for all rational numbers r. Hence, for all rational numbers r we have

$$0 \leqq (rx + y, rx + y) = r^2(x, x) + 2r(x, y) + (y, y),$$

and so this implies the Cauchy-Schwartz inequality $|(x, y)| \leqq \|x\| \cdot \|y\|$. This in turn leads directly to the triangle inequality.

Also solved by R. A. Avelsgaard, G. Baron & W. Imrich (Austria), Robert Bowen, Robert Cohen, J. A. Dyer, E. J. Eckert, M. G. Greening (Australia), Harry Guess, D. C. Kay, E. S. Langford, Michael Lempel, A. E. Livingston, Donald Pilling, N. R. Riesenberg, D. A. Robinson, V. L. N. Sarma (India), Robin Sibson (England), S. J. Sidney, Richard Sinkhorn, John Stout, W. C. Waterhouse, R. J. Weinacht, J. S. W. Wong, D. A. Zave, and the proposer.

A number of erroneous solutions were received for this problem, all of which pursued, in one form or another, the same circular argument proceeding by contra position to assume (i) $\|x+y\|>\|x\|+\|y\|$, replacing y by $-y$ and using also (ii) $\|x-y\|>\|x\|+\|y\|$ in order to obtain a final contradiction. As a matter of fact, the parallelogram law and (i) yield directly the triangle inequality for x and $-y$, but clearly (i) is not thereby contradicted.

Sum of Ratios of Sines

5265 [1965, 194]. *Proposed by E. Ehrhart, Strasbourg, France*

For which integral values of a, b, c is the following sum defined and equal to an integer:

$$\sum_{n=1}^{a-1} \frac{\sin(bn\pi/a)}{\sin(cn\pi/a)}?$$

Note by Al Somayajulu, University of Rochester. The solution to this problem is given in full by N. J. Fine in connection with problem E 1683 [1965, 319]. In particular, see the Editor's résumé for E 1683 on p. 321.

Also solved by Sahib Ram Mandan (India), and by the proposer.

Iterates of an Increasing Convex Function

5266 [1965, 194]. *Proposed by J. D. Pryce, University of Newcastle, England*

Let F be a positive nondecreasing, convex function on $[0, \infty)$ with $F(0)=0$ and with $y=x-\alpha$ as an asymptote, $(\alpha>0)$. It is easily seen that the function $H(x)=\frac{1}{2}F(F(2x))$ again has $y=x-\alpha$ as asymptote. Prove $H(x)\leqq F(x)$ for all x.

Hence prove that if we define a sequence $\{F_n\}$ by

$$F_1 = F, \qquad F_n(x) = \tfrac{1}{2}F_{n-1}(F_{n-1}(2x)),$$

then every F_n is a positive nondecreasing convex function and $F_n(x) \to \max(0, x-\alpha)$ uniformly on $[0, \infty)$.

Solution—Composition of solutions submitted by D. Ž. Djoković, University of Belgrade, Yugoslavia, A. E. Livingston, University of Alberta, and S. J. Sidney, Harvard University. The conditions of the problem clearly imply that $F(x+h) \leqq F(x)+h$ for all $x \geqq 0$, $h \geqq 0$. Thus it follows that $F(2x) \leqq F(x)+x$ and

$$H(x) = \tfrac{1}{2}F(F(2x)) \leqq \tfrac{1}{2}F(x + F(x)) \leqq F(x). \tag{1}$$

That $y=x-\alpha$ is an asymptote of $H(x)$ follows now from $H(x) \leqq F(x)$ and

$$H(x) = \tfrac{1}{2}F(F(2x)) \geqq \tfrac{1}{2}F(2x - \alpha) \geqq \tfrac{1}{2}(2x - \alpha - \alpha) = x - \alpha.$$

The convexity of $H(x)$ follows from the convexity of $F(x)$; we have

$$\begin{aligned} H(\alpha x + (1 - \alpha)y) &= \tfrac{1}{2}F(F(2\alpha x + 2(1 - \alpha)y)) \leqq \tfrac{1}{2}F(\alpha F(2x) + (1 - \alpha)F(2y)) \\ &\leqq \tfrac{1}{2}\alpha F(F(2x)) + \tfrac{1}{2}(1 - \alpha)F(F(2y)) = \alpha H(x) + (1 - \alpha)H(y). \end{aligned}$$

Thus $H(x)$ and each $F_n(x)$ is a convex nondecreasing function, positive for $x>0$, and $\max(0, x-\alpha) \leqq F_n(x) \leqq F_{n-1}(x)$ for all $x \geqq 0$.

It follows that $F_n(x)$ converges to a nondecreasing, convex continuous function $G(x)$ with asymptote $y=x-\alpha$, $x \geqq 0$. Moreover the convergence is uniform on every finite interval $[0, x_0]$. We have therefore

$$\max(0, x - \alpha) \leqq G(x) = \lim_{n\to\infty} F_{n+1}(x) = \lim_{n\to\infty} \tfrac{1}{2}F_n(F_n(2x)) = \tfrac{1}{2}G(G(2x)).$$

But with the same reasoning leading to (1) above,

$$\tfrac{1}{2}G(G(2x)) \leqq \tfrac{1}{2}G(x + G(x)) \leqq \tfrac{1}{2}(G(x) + G(x)) = G(x)$$

and it follows that $G(2x)=G(x)+x$ when $G(x)$ is strictly increasing. This implies that either $G(x)=0$ or $G(t)$ is linear for $t \geqq x$, i.e. $G(t)=t-\alpha$, $t \geqq x$. It follows now that $G(x)=\max(0, x-\alpha)$ and the uniformity on $[0, \infty)$ is a consequence of the fact that $F_n(\alpha)=\sup_{x\geqq 0}[F_n(x)-\max(0, x-\alpha)]$ and $F_n(\alpha) \downarrow 0$.

Editorial Note. In his solution, Livingston also proved that $y=x-\alpha$ is an asymptote of $H(x)$ whenever it is an asymptote of $F(x)$ regardless of the monotone or convex nature of $F(x)$.

Primitive Permutation Groups

5268 [1965, 194]. *Proposed by V. H. Keiser, University of Colorado*

(a) A transitive nilpotent permutation group which is not of prime degree is imprimitive.

(b) A primitive permutation group which is not of prime degree has no center.

Solution by Gomer Thomas, University of Illinois. (a) If a transitive nilpotent permutation group is primitive, then the stabilizer of every symbol must be maximal, and hence normal of prime index. Hence the degree of the group must be prime, since it is equal to the index of any stabilizer. This is equivalent to the desired result.

(b) Let G be primitive and not of prime degree, and let Z be the center of G. Since Z is normal in G, either $Z=\{1\}$ or Z is transitive. If Z were not $\{1\}$, then by part (a) Z would have to be imprimitive. But any block for Z is also a block for G, so the imprimitive nature of Z would imply the imprimitive nature of G, contrary to hypothesis. It follows that $Z=\{1\}$.

Also solved by M. G. Greening (Australia), C. C. Lindner, C. R. MacCluer, W. C. Waterhouse, and the proposer.

Separation in First Countable Spaces

5269 [1965, 194]. *Proposed by Michael Gemignani, University of Notre Dame*

Prove that any first countable topological space in which every compact subset is closed is T_2.

Solution by S. P. Franklin, University of Florida. In a first countable but non-Hausdorff space one easily constructs a sequence converging to two distinct points. The range of this sequence together with one of its limit points forms a nonclosed compact subset.

Also solved by J. S. Alin, D. R. Anderson, D. R. Andrew, R. O. Atkinson, S. Baron, Ralph Bennett, Robert Bowen, George Cain, Jr., A. G. Dotson, Jr., James Joseph, B. G. Klein, Anne B. Koehler, Shu-Chung Koo, A. E. Livingston, Peter Loeb, M. D. Mavinkurve (India), M. R. Meck, J. C. Morgan II, S. A. Naimpally, F. J. Papp, Jr., L. K. Paul, Jr., W. J. Perwin, R. L. Plunkett, S. M. Robinson, V. L. N. Sarma (India), P. S. Schnare, G. F. Schumm, S. J. Sidney, David A. Smith, W. R. Smythe, Jr., Al Somayajulu, J. C. Wenger, Margaret R. Wiscomb, D. A. Zave, and the proposer.

Several contributors pointed out that the result is known and may be found in the following papers in this MONTHLY:

(1) A. J. Insel, *A note on the Hausdorff separation property in first countable spaces*, 72(1965) 289.

(2) Norman Levine, *When are closed and compact equivalent?*, 72(1965) 41.

(3) Edwin Halfar, *A note on Hausdorff separation*, 68(1961) 164.

Any historian knows that George Washington was born in $[1000\sqrt{3}]$.

JACK C. ABAD, City College of San Francisco

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University
COLLABORATING EDITORS: K. O. MAY, University of California, Berkeley, and
E. P. VANCE, Oberlin College

Materials intended for review should be sent directly as follows: Books: R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457. Programmed Materials: K. O. May, University of California, Berkeley, Calif. 94704. Films: E. P. Vance, Oberlin College, Oberlin, Ohio 44074.

Introduction to Mathematical Logic. By Elliott Mendelson. Van Nostrand, Princeton, N. J., 1964. 300 pp. $7.25.

This is a short but substantial introduction to some of the important ideas of mathematical logic. The author says that the text will be used in departments of mathematics and philosophy at the junior, senior and graduate level but, like most authors, he is optimistic. A graduate student in mathematics will probably get a great deal from the book—a junior in philosophy, very little. Sometimes the pace at which the author plunges into difficult ideas is likely to leave the reader wondering what has happened. A rather more lengthy introduction and discussion of the elementary aspects whenever a new topic is introduced would help the reader considerably.

The first chapter begins with the elementary ideas of the propositional calculus. The main result proved is the Deduction Theorem. The second chapter deals with the first order predicate calculus. Henkin's proof of Gödel's Completeness Theorem is given as well as an introduction to model theory. Chapter 3 is devoted mainly to a proof of Gödel's Incompleteness Theorem for a particular formal theory.

These first three chapters form a good short course in mathematical logic for mathematically mature students.

The remaining two chapters take up axiomatic set theory and the theory of algorithms. In Chapter 4, a version of von Neumann's axioms for set theory is used and the development carried through up to transfinite induction and the axiom of choice. Chapter 5 is, for a book of this level, a very good and quite detailed study of the notion of effective computability. The author develops the theory of Markov algorithms to the point of proving the equivalence of Markov-computability and recursiveness. Turing machines and Herbrand-Gödel computability are also shown to lead to the same class of "effectively computable" functions. This chapter concludes with a brief discussion of some undecidable problems.

There are a large number of minor misprints and errors, most of which the reviewer would not have noticed if his attention had not been drawn to them by a communication from the author. All in all, however, this book is a useful addition to the books bringing very recent work in mathematical logic down to the undergraduate or first-year graduate level. The only major comment the re-

viewer has to make has already been made above—if the book is to be used by students without any prior knowledge of mathematical logic, a considerable knowledge of other abstract mathematics is very desirable.

TREVOR EVANS, Emory University

Mathematisches Wörterbuch. Edited by Joseph Naas and H. L. Schmid. Teubner, Stuttgart, 1961, and Pergamon, New York, 1962. Two volumes, viii+1043 and iv+952 pp. $110.

This encyclopedic dictionary is the work of over a hundred collaborators associated with the Institute of Pure Mathematics of the German Academy of Science of Berlin. It includes brief definitions, surveys of broad topics, concepts, applications, methods, and tendencies of current research, over 400 biographies, bibliographies, and a sixteen page list of symbols classified under 82 subject headings. Extensive cross references facilitate tracing related ideas. The coverage is extensive, intensive, and up to date. The level is from college undergraduate to post-doctoral. It is a mathematician's dictionary. The printing and binding are excellent. There is no seriously competitive collection of mathematical information in any language. Even for someone whose knowledge of German is meager the quickest way to information on a mathematical topic might be to look up the German word in an English-German dictionary and then turn to Naas and Schmid. The price is staggering, even after the 20% professional discount offered by Pergamon. Nevertheless, no mathematical library should be without it, and every research worker ought to have quick access to a copy. Obviously there should be such a work in English, but the need for adequate mathematical reference books seems not yet to be recognized.

KENNETH O. MAY, University of California, Berkeley

A Geometric Introduction to Linear Algebra. By Daniel Pedoe. Wiley, New York, 1963. xi+224 pp., $5.95.

The title of this book should be taken literally. It is an *introduction to* linear algebra and not a textbook on that subject. Moreover, the introduction is *geometric* in that it is firmly based on geometric concepts which are developed in the first four chapters. The topics covered are analytic geometry in the plane (points, lines and direction ratios only); vectors in the plane; planes and lines in 3-space, vectors in 3-space; regular systems of linear equations with emphasis on elementary row transformations of the matrix of the system; vector spaces, linear dependence, bases, subspaces and inner products, (the emphasis here is on subspaces of the space of n-tuples of real numbers); matrix algebra, rank, general systems of linear equations, and determinants; a final short chapter deals with linear mappings and their associated matrices.

In the reviewer's opinion the choice of subject matter makes good sense pedagogically and the book should be useful to at least two groups of students. For the mathematics major it provides a good background in the more concrete and computational aspects of matrix algebra as well as some insight into the

generalizations of familiar geometric concepts that are possible in space of n dimensions. On completion of this book a student would be well equipped for a sophisticated course in linear algebra. On the other hand, for the increasing number of students whose major field is not mathematics, but who require some knowledge of matrix algebra and linear systems, an excellent terminal course could be based on this book, although some supplementation would be desirable. The most obvious lack is a systematic discussion of transformation of coordinates from one Cartesian system to another. In particular rotation of axes is not treated either in the plane, in 3-space or in R_n. This seems a rather serious omission in a book of this title. Reduction of quadratic forms could also well be included in a course at this level. A course should be generated by a text book, however, not coincide with it, and with this viewpoint the book should have a wide appeal.

D. C. MURDOCH, University of British Columbia

Point Set Topology. By S. A. Gaal. Pure and Applied Mathematics, Vol. XVI. Academic Press, New York and London, 1964. xi+317 pp. $9.75.

Chapter titles and numbers of pages in each are: Introduction to Set Theory (20), Topological Spaces (56), Separation Properties (50), Compactness and Uniformization (48), Continuity (79), and Theory of Convergence (55). Sections are frequently short, include examples, and are followed by exercises, some of which are routine and some introduce new concepts. Hints for their solutions are very often provided; this means that the student has less chance to develop facility and inventiveness in discovering proofs. There are several paragraphs of notes at the end of each chapter which mention un-investigated areas and unsolved problems. These are followed by an extensive list of references. There is a list of symbols and a comprehensive index.

In his preface, the author states that this book is for beginning graduate students and advanced undergraduates, and that it presents point set topology not only as an end in itself, but also as a related discipline to the proper understanding of various branches of analysis and geometry. He also says that the book contains enough material for a two semester course; this is unquestionably true. He treats the material with great care and unusual depth. Compactness, for example, appears with these prefixes: bi, countable, countably para, hypo, local, local (m, n), meta, one-point, para, pre, quasi, rim, sigma, sequential, and strongly para. There is a really exhaustive study of separation properties and their interrelationships.

A one or two semester undergraduate course could probably be constructed by careful choice of topics; the reviewer feels that better texts for this audience already exist. The book would be excellent as a reference work, and probably quite satisfactory as a graduate text, depending on the instructor's prejudices.

No serious misprints were found; the ones on page 12, for example, are quickly corrected.

D. S. RAY, Bucknell University

A Course of Higher Mathematics, Volume III-2. By V. I. Smirnov. Complex Variables—Special Functions. Translated by D. E. Brown, edited by I. N. Sneddon. Addison-Wesley, Reading, Mass.; Palo Alto; London, 1964. 700 pp. $15.00.

The volume is part of an extensive and comprehensive course in mathematical analysis, intended for applied mathematicians, physicists and theoretical engineers. The entire five volume work culminates in the differential, integral and functional equations of mathematical physics, and all its parts are coordinated and edited with this ultimate end in mind. The volume reviewed here deals with the theory of analytic functions, linear differential equations in the complex domain and the special functions which occur in the series developments of mathematical physics. The formal and computational part of the theory of complex variables is clearly and carefully treated and illustrated by numerous examples taken from fluid dynamics, elasticity, electromagnetic theory, and quantum theory. Frequently, questions of purely mathematical interest are only very briefly mentioned. Thus, large discussions are devoted to the theory of special conformal mappings, but the basic Riemann mapping theorem is only quoted but not proved. This limitation is clearly needed to leave space for the very detailed study of the constructive part of complex analysis. The book could be used to advantage in courses on complex variable theory for engineers, or on the theory of special functions, etc., but could not serve as the basis of an advanced course in complex variables for mathematicians. It is particularly useful as a reference work, preferably within the framework of the whole set, which is a valuable repository of information in modern applied analysis.

A word should be said about the English translation. One feels that the translator is no mathematician and that the editor has been somewhat careless. For example, many additional topics are mentioned briefly at the end of the sections, and the reader is referred to books and papers for more penetrating study. In the original, these literary quotations refer naturally to Russian publications. In the translation, the editor would have performed a service by replacing such references by more available papers and books in English. A student who has learned the theory of complex variables from the translation would know the beautiful and important theorem of Sokhotski. He could never communicate, however, with a fellow student who used a western textbook where the theorem is ascribed to Casorati and Weierstrass. I do not enter upon the delicate question of priority but point out only the difficulty in looking up or quoting such theorems. Next, we consider the retransliteration of western names from the Russian. If I had not known an important formula of Goursat, I would have believed that an ingenious man by the name of Hurse discovered a new representation for biharmonic functions. I suspect that there is no man by the name of Leber and that his integral belongs to Lebesgue. Who are Eiry, Mizes and Gilbert? (Airy, Mises, Hilbert?) Could a student ask in the library for the

"News of the Bavarian Academy"? A little more care in editing would have increased considerably the value of the translation. Let the reader be warned to be skeptical in the use of all references and quotations.

M. M. SCHIFFER, Stanford University

Tutorial Texts and Problem Collections in Mathematics, Volume I. Elementary Inequalities. By D. S. Mitrinović (in collaboration with E. S. Barnes, D. C. B. Marsh, and J. R. M. Radok). Noordhoff, Gröningen, 1964. ix+150 pp. $5.75.

This monograph contains an extensive specialized selection of problems. The inequalities range from quite easy to very difficult; every problem solver ought find some to his interest. Many of the better known inequalities are accompanied by proofs, but a few of the basic ones are assumed known to the reader. There is some attempt to bring order by the separation of the book into sections; this may well mean that it can be used occasionally as a reference. Section 10, Inequalities in the Complex Domain, is regrettably brief for my tastes.

As far as independent study is concerned, let the beginner beware; the student may well need help and guidance. One of the problems left for the student is either to develop some general theory or to go elsewhere for it. The proofs are often more ingenious than instructive, suggestive geometrical sketches are scarce, references are few in number, and discussions relating the problems to other material (such as their sources) are virtually non-existent. It seems to me that there are several other books now available which are superior for providing insight into methods, although less extensive. I worry especially about the isolation of the study of inequalities from their occurrence.

R. G. BUSCHMAN, SUNY at Buffalo

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor C. J. Pipes, Southern Methodist University, represented the Association at the inauguration of James M. Moudy as Chancellor of Texas Christian University on November 19, 1965.

Professor Malcolm Pownall, Colgate University, represented the Association at the inauguration of G. B. Dearing as President of the State University of New York at Binghamton on September 25, 1965.

Professor G. G. Roberts, Berea College, has been awarded the Seabury Award for Excellence in Teaching.

University of Alabama: Dr. J. M. Horner, General Motors Research Laboratories, Warren, Michigan, and Dr. W. J. Gray, University of Alabama, have been appointed Assistant Professors; Associate Professor J. L. Howell has been promoted to Professor; Professor F. A. Lewis retired in September 1965 with the title of Professor Emeritus.

University of Alaska: Assistant Professor Thomas Head, Iowa State University, has been appointed Professor; Assistant Professor P. A. Van Veldhuizen has been promoted to Associate Professor.

University of Alberta, Calgary: Professor W. A. Al-Salam, Texas Technological College, has been appointed Associate Professor; Mrs. N. A. Al-Salam, Texas Technological College, has been appointed Assistant Professor.

Auburn University: Professor L. P. Burton has been promoted to Head Professor; Associate Professor Emilie V. Haynsworth has been promoted to Professor; Assistant Professors C. E. Robinson and Margaret Baskervill have been promoted to Associate Professors.

Brooklyn College: Assistant Professors Margaret Y. Woodbridge and L. S. Kennison have been promoted to Associate Professors.

California State Polytechnic College: Mr. D. R. Stevens, Upland High School, Upland, California, and Mr. N. R. Townsend, Wisconsin State College, Stevens Point, have been appointed Assistant Professors; Associate Professor W. O. Buschman has been promoted to Professor.

University of California, Irvine: Professor G. K. Kalisch, University of Minnesota, has been appointed Professor; Professor Emeritus Einar Hille, Yale University, has been appointed Senior Lecturer; Dr. Robert Maltz, University of Paris, France, and Dr. Bernard Russo, University of California at Los Angeles, have been appointed Assistant Professors.

University of California, Riverside: Dr. J. L. Denny, Jr., Indiana University, Dr. J. E. de Pillis, San Francisco State College, and Dr. Le Baron O. Ferguson, University of Washington, have been appointed Assistant Professors.

University of California, Santa Barbara: Professor Ky Fan, Northwestern University, has been appointed Professor; Dr. S. C. Bachmuth, Fairleigh Dickinson University, has been appointed Assistant Professor; Assistant Professor R. C. Thompson has been promoted to Associate Professor; Drs. David Outcalt and Melvin Rosenfeld have been promoted to Assistant Professors.

Chatham College: Assistant Professor D. H. Trahan, University of Pittsburgh, has been appointed Assistant Professor and Acting Chairman of the Department of Mathematics; Associate Professor W. A. Beck is on leave as Fulbright Lecturer at Al-Hikma University in Baghdad, Iraq.

City College of New York: Dr. William Wernick, Evander Childs High School, Bronx, New York, has been appointed Associate Professor; Assistant Professor Sidney Penner, State University of New York at Buffalo, and Dr. N. R. Wagner, Massachusetts Institute of Technology, have been appointed Assistant Professors.

Colorado State University: Dr. H. H. Frisinger has been promoted to Assistant Professor; Assistant Professor K. J. Whitcomb has been promoted to Associate Professor; Professor A. G. Clark retired in June 1965 with the title of Professor Emeritus.

Cornell University: Dr. S. A. Levin, University of California at Berkeley, has been appointed Assistant Professor; Associate Professor Anil Nerode has been promoted to Professor.

University of Delaware: Dr. R. H. Wenger, Michigan State University, has been

appointed Assistant Professor; Professor Russell Remage, Jr. has been appointed Chairman of the Department of Mathematics.

Denison University: Professor Chosaburo Kato retired with the title of Professor Emeritus; Professor Marion Wetzel has been named to the Benjamin Barney Chair of Mathematics; Professor Andrew Sterrett has been named Chairman of the Department of Mathematics; and Associate Professor W. N. Prentice has been named Director of the Computer Center.

East Carolina College: Drs. D. F. Bailey, Vanderbilt University, and Katye O. Sowell, Florida State University, have been appointed Assistant Professors; Mrs. Stella M. Daugherty has been promoted to Assistant Professor; Assistant Professor C. A. Webber, Jr. has been promoted to Associate Professor; Associate Professor Louise L. Williams has been promoted to Professor.

Emory University: Mr. J. P. Downes, Columbia University, Dr. D. A. Ford, Marshall Space Flight Center, and Dr. O. T. Nelson, Vanderbilt University, have been appointed Assistant Professors; Assistant Professor B. K. Youse has been promoted to Associate Professor.

Georgia Institute of Technology: Assistant Professor S. H. Coleman, University of Wisconsin, has been appointed Associate Professor; Assistant Professor M. Z. Nashed has been promoted to Associate Professor; Associate Professor G. C. Caldwell has been promoted to Professor.

University of Hawaii: Dr. W. J. Leahey, University of Illinois, has been appointed Assistant Professor; Assistant Professor Frances E. Davis has been promoted to Associate Professor; Mrs. Ruth E. M. Wong has been promoted to Assistant Professor.

Hofstra University: Assistant Professor L. E. Sigler, Hunter College, has been appointed Assistant Professor; Associate Professor G. B. Charlesworth has been appointed Acting Chairman.

University of Houston: Professor P. D. Hill, Emory University, has been appointed Professor; Associate Professor J. N. Younglove, University of Missouri, has been appointed Associate Professor; Assistant Professor R. D. Sinkhorn has been promoted to Associate Professor; Assistant Professor D. R. Traylor has been appointed Chairman of the Mathematics Department.

University of Illinois: Professor Paul Erdös, Israel Institute, has been appointed G. A. Miller Visiting Professor; Professor J. L. Selfridge, Pennsylvania State University, has been appointed Professor; Associate Professor H. A. Osborn has been promoted to Professor; Assistant Professors R. L. Bishop and P. G. Braunfeld have been promoted to Associate Professors; Mrs. Lucretia S. Levy, Dr. T. G. McLaughlin, Mr. D. R. Sherbert, and Mr. Frederick Hoffman have been promoted to Assistant Professors; Associate Professor Echo D. Pepper retired on September 1, 1965, with the title of Associate Professor Emerita.

Indiana State College: Mr. J. A. Peters, Ohio State University, has been appointed Associate Professor; Mrs. Emma Lou Somers, Shippensburg State College, has been appointed Assistant Professor.

University of Maryland: Dr. H. K. Martens, Institutt for Matematiske Fag, Norway, has been appointed Visiting Associate Professor; Assistant Professor C. H. Cook, University of Oklahoma, and Dr. S. L. Gulick, III, University of Pennsylvania, have been appointed Assistant Professors; Associate Professor B. L. Reinhart has been promoted to Professor; Assistant Professor R. S. Freeman has been promoted to Associate Professor.

University of Massachusetts: Professors Y. W. Chen, Wayne State University, Haskell Cohen, Louisiana State University, Wladimir Seidel, Wayne State University, and Associate Professor D. J. Foulis, University of Florida, have been appointed Professors;

Associate Professor Berthold Schweizer, University of Arizona, and Assistant Professor W. W. Comfort, University of Rochester, have been appointed Associate Professors; Dr. D. E. Catlin, University of Florida, and Dr. Murray Eisenberg, Wesleyan University, have been appointed Assistant Professors; Professor A. E. Andersen retired on September 30, 1965 with the title of Professor Emeritus.

University of Miami: Professor P. R. Halmos, University of Michigan, has been appointed Visiting Professor; Assistant Professor E. A. Nordgren, University of New Hampshire, has been appointed Visiting Assistant Professor; Associate Professor R. W. Bagley has been promoted to Professor; Assistant Professor Edwin Duda has been promoted to Associate Professor.

University of New Hampshire: Dr. Samuel Shore, Pennsylvania State University, has been appointed Assistant Professor; Professor M. E. Munroe, Chairman of the Mathematics Department, is on leave under a Fulbright Grant at the University of Cairo; Associate Professor E. H. Batho has been appointed Acting Chairman of the Department of Mathematics.

Nicholls State College: Mr. Larry Haw, Morrow High School, Morrow, Louisiana, has been appointed Assistant Professor; Assistant Professor Loraine M. Cook has been promoted to Associate Professor.

University of Notre Dame: Associate Professor R. R. Otter has been promoted to Professor; Assistant Professor C. R. Riehm has been promoted to Associate Professor; Associate Professor T. E. Stewart has been promoted to Professor and appointed Associate Vice President of Academic Affairs.

Oakland University: Associate Professor G. P. Johnson, University of the South, has been appointed Professor and Chairman of the Mathematics Department; Assistant Professor D. G. Malm has been promoted to Associate Professor.

Sacramento State College: Mr. C. H. Tjoelker, Oregon State University, has been appointed Assistant Professor; Associate Professor G. R. Glabe has been promoted to Professor.

Shippensburg State College: Mr. J. S. Mowbray, Lehigh University, has been appointed Assistant Professor; Dr. P. F. Cauffman has been promoted to Dean of Teacher Education.

South Dakota School of Mines and Technology: Professor Harold Heckart, Illinois College, has been appointed Associate Professor; Mr. Wayne Walther, Univac, St. Paul, Minnesota, has been appointed Assistant Professor; Professor D. C. Benson has been appointed Chairman of the Mathematics Department; Associate Professor C. A. Grimm has been promoted to Professor.

Southern Illinois University: Associate Professor Neal Foland, Kansas State University, and Assistant Professor R. A. Moore, Pennsylvania State University, have been appointed Associate Professors; Mr. Zamir Bavel has been promoted to Assistant Professor; Dr. Carl Townsend, Washington State University, has been appointed Assistant Professor.

Syracuse University: Associate Professors P. T. Church and S. C. Moy have been promoted to Professors; Assistant Professors John Lindberg and J. D. Reid have been promoted to Associate Professors.

Texas Christian University: Professor O. H. Hamilton, Oklahoma State University, has been appointed Professor; Assistant Professor C. R. Deeter has been promoted to Associate Professor; Professor C. R. Sherer retired on September 1, 1965 with the title of Professor Emeritus.

Texas Technological College: Professors A. R. Amir-Moéz, Clarkson College of Technology, and J. W. Ault, U. S. Air Force Academy, have been appointed Professors; Assistant Professors E. F. Steiner, University of New Mexico, and J. T. White, Univer-

sity of Kansas, have been appointed Associate Professors; Assistant Professors T. A. Atchison, A. A. Gioia, and S. K. Hildebrand have been promoted to Associate Professors; Mr. H. L. Gray, University of Texas, has been appointed Assistant Professor.

Texas Western College: Associate Professor D. L. Boyer, University of Idaho, has been appointed Professor; Assistant Professor E. L. Allgower, Sacramento State College, has been appointed Associate Professor.

Trinity University: Associate Professor D. L. George, Charlotte College, has been appointed Associate Professor; Mr. Gerald Smetzer has been promoted to Assistant Professor.

University of Virginia: Assistant Professor K. O. Leland, Ohio State University, has been appointed Acting Assistant Professor; Associate Professor Marvin Rosenblum has been promoted to Professor; Dr. J. W. England has been promoted to Assistant Professor.

Washington State University: Associate Professors W. E. Barnes and C. T. Long have been promoted to Professors.

West Virginia University: Assistant Professors C. N. Cochran and H. W. Gould have been promoted to Associate Professors; Professor I. D. Peters has been appointed Acting Chairman of the Mathematics Department; Professors H. A. Davis and C. H. Vehse retired with the title of Professor Emeritus.

Western Washington State College: Associate Professor Byron McCandless, Rutgers, The State University, has been appointed Professor; Assistant Professor Andre Yandl, Seattle University, has been appointed Associate Professor; Assistant Professor J. R. Reay has been promoted to Associate Professor.

Wisconsin State University, Whitewater: Professor G. H. Miller, Parsons College, has been appointed Associate Professor; Mr. R. L. Boehning, University of Wisconsin, Green Bay Center, and Mr. Ronald Dettmers, University of Notre Dame, have been appointed Assistant Professors; Associate Professor L. R. Tappan, Beadle State College, has been appointed Associate Professor.

University of Wyoming: Professor Emeritus B. G. Clark, Vanderbilt University, has been appointed Visiting Professor; Dr. J. A. Jensen, Iowa State University, Dr. J. S. Rue, Iowa State University, and Dr. R. A. Stoltenberg, New Mexico State University, have been appointed Assistant Professors.

Professor L. C. Bagby, Detroit Institute of Technology, has retired to become Vice President of International Casting Corporation, Detroit, Michigan.

Associate Professor Truman Botts, University of Virginia, is on a year's leave of absence to serve as Executive Director of the National Academy of Science Committee on Support of Research in the Mathematical Sciences with headquarters at Columbia University.

Mr. Phillip Briggs, East Texas State College, has been appointed Assistant Professor at East Central State College.

Dr. D. L. Bruyr, Kansas State Teachers College, has returned from a leave of absence spent at Oklahoma State University and has been appointed Associate Professor.

Professor G. R. Bushyager, Morningside College, retired in June 1965.

Professor E. A. Cameron, University of North Carolina, has been awarded a National Science Foundation Science Faculty Fellowship for 1965–66 and is spending the year at Harvard University.

Professor K. H. Carlson, Valparaiso University, has been appointed Visiting Professor at Brown University.

Professor D. E. Christie, Bowdoin College, has been promoted to Wing Professor of Mathematics.

Dr. L. H. Coon, Ohio State University, has been appointed Associate Professor at Eastern Illinois University.

Mr. K. H. Crawford, Berea College, has returned from a year's study at Florida State University and has been appointed Assistant Professor.

Assistant Professor H. K. Crowder, Case Institute of Technology, has been appointed Associate Professor at Cleveland State University.

Associate Professor Joyce C. Cundiff, State University of New York at Fredonia, has been appointed Associate Professor at Midwestern University.

Assistant Professor David Dautenhahn, University of Missouri at Rolla, has been appointed Associate Professor at Missouri Valley College.

Dr. J. E. Diem, Purdue University, has been appointed Assistant Professor at Tulane University.

Associate Professor E. D. Dixon, West Georgia College, has been appointed Associate Professor at Tennessee Technological University.

Assistant Professor Forrest Dristy, Florida Presbyterian College, has been promoted to Associate Professor.

Assistant Professor D. J. Ewy, Fresno State College, has been promoted to Associate Professor.

Dr. J. R. Hanson, Virginia Polytechnic Institute, has been appointed Assistant Professor at Madison College.

Professor R. C. Huffer, Head of the Mathematics Department, Tougaloo College, has also been appointed Acting Academic Dean.

Rev. Ronald King, St. Louis University, has been appointed Assistant Professor at Siena College.

Dr. C. F. Kossack, Laboratory for Computer Sciences, Dallas, Texas, has been appointed Professor of Statistics and Head of the Department of Statistics at the University of Georgia.

Mr. G. E. Lindamood, University of Maryland, has been appointed Instructor in the Computer Science Center.

Mr. D. H. McInnis, Southwest Missouri State College, has been promoted to Assistant Professor.

Assistant Professor Maurice Machover, Fairleigh Dickinson University, has been promoted to Associate Professor.

Mr. Bernard Martin-Williams, Christ Church School, Greenville, South Carolina, has been appointed Associate Professor at West Georgia College.

Mr. H. L. Minton, University of Texas, has been appointed Assistant Professor at Memphis State University.

Associate Professor D. H. Moore, California State Polytechnic College, has been promoted to Professor.

Associate Professor W. O. Portmann, Arizona State University, has been appointed Professor and Chairman of the Mathematics Department at Wilson College.

Assistant Professor A. A. Ramsay, Brandeis University, has been appointed Assistant Professor at the University of Rochester.

Associate Professor Herbert Rebassoo, Luther College, has been promoted to Professor.

Associate Professor K. W. Reed, Texas Southern University, has been appointed Visiting Assistant Professor at Rice University.

Dr. Jimmy Rice, Fort Hays Kansas State College, has been promoted to Professor.

Assistant Professor E. J. Schweppe, University of Maryland, has been promoted to Associate Professor in the Computer Science Center.

Professor S. M. Selby, University of Akron, has been awarded a Distinguished Professorship.

Assistant Professor Richard Shermoen, North Dakota State University, has been promoted to Associate Professor.

Dr. R. L. Shively, Oak Ridge National Laboratory, Oak Ridge, Tennessee, has been appointed Professor and Chairman of the Mathematics Department at Lake Forest College.

Mr. J. T. Smith, Martin Company, Denver, Colorado, has been appointed Assistant Professor at Northern Montana College.

Dr. L. C. Sulski, University of Sussex, England, has been appointed Assistant Professor at Holy Cross College.

Dr. J. F. Traub, Bell Telephone Laboratories, Murray Hill, New Jersey, has been appointed Visiting Associate Professor in the Computer Science Department at Stanford University for the spring of 1966.

Assistant Professor R. W. Tucker, Humboldt State College, has been promoted to Associate Professor.

Perkins Professor J. L. Walsh, Harvard University, retired February 1966 with the title of Perkins Professor of Mathematics Emeritus and has been appointed Professor at the University of Maryland.

Mr. Henry Wellenzohn, Niagara University, has been promoted to Assistant Professor.

Professor R. E. Wheeler, Howard College, has been promoted to Chairman of the Division of Natural Sciences and Head of the Mathematics Department.

Dr. E. T. Wooldridge, Florence State College, has been promoted to Professor.

Mr. J. E. Wynn, Utah State University, has been appointed Lecturer at California State Polytechnic College.

Mr. William Betz, Rochester Public Schools, Rochester, New York, died on September 7, 1965. He was a charter member of the Association.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

NOVEMBER MEETING OF THE MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA SECTION

The annual Fall meeting of the Maryland-District of Columbia-Virginia Section of the MAA was held at Goucher College, Towson, Maryland, on November 21, 1964. Dr. Daniel Shanks, Chairman of the section, presided. There were 98 persons in attendance, including 92 members of the Association.

The following program was presented:

1. *On the number of non-equivalent types of triple-systems of order* $N \geqq 15$, by V. Bohun-Chudyniv, Morgan State College, Baltimore, Maryland.

In 1893 E. Moore proved the existence of at least two nonequivalent types of triple-systems of order greater than 13. Zulauf in 1897 established the existence of two non-equivalent types of triple-systems of order 13. All examples constructed by Kirkman in 1847, Reiss in 1889, De Vries

in 1889 and 1894, belong to the first type, while those constructed by Netto's method are of the second type. In 1850 A. Cayley constructed for the solution of Kirkman's problem two examples of triple-systems of the 15-th order. Many other papers have been published concerning triple-systems of orders higher than 15. In the present paper the author, by using modular triple-systems and block-operators, establishes the existence of at least 8 nonequivalent types of triple-systems of $N \geqq 15$ order, and enumerates the triple-systems of the first type.

2. *Remarks on relaxation methods and computers*, by M. N. McAllister, Towson, Maryland.

The author presents a brief description of the fundamental ideas involved in the use of the relaxation methods in solving the problem $Px=y$, where y is given and P is any operator. Reasons are given why the relaxation methods have never been used in relation to computers. To overcome this inherent difficulty is a new field of research. Very good results have been obtained by combining iteration with relaxation methods. This is illustrated by presenting the Conte-Dames method for solving the biharmonic equation $\nabla^4 w=0$. Some problems, not yet resolved, are pointed out.

3. *Sifting properties of singularity functions in the solution of linear differential equations*, by F. Marion Clarke-Carroll, Westinghouse Electric Corporation, Baltimore, Maryland.

Taking note of the computational clumsiness engendered by nonzero initial conditions of a linear differential system, the speaker investigates a method by which a homogeneous linear equation with arbitrary initial conditions can be replaced by a corresponding linear differential equation with zero initial conditions and the same solution. To this end certain singularity functions and their Laplace transforms are employed and certain sifting properties which make them useful in this context are demonstrated.

4. *Mathematical Theory in Optimization Techniques*, by Walker H. Land, Jr., I.B.M., Bethesda, Maryland.

This paper discusses the necessary theoretical background in the mathematical techniques which are utilized in trajectory optimization. Included are derivations, from fundamental principles, of the following topics: (a) The Calculus of Variations—necessary and sufficiency conditions; (b) relationship between the Maximum Principle of Pontryagin and the Weierstrass Excess Function; (c) variational concepts in terms of the Hamiltonian; (d) dynamic programming and the Calculus of Variations; (e) method of Steepest Ascent or Method of Gradients.

5. *Another look at the so-called Andoyer-Lambert Distance Formulas*, by Paul D. Thomas, Coast and Geodetic Survey, Washington, D. C.

These approximation formulae, of first order in the ellipsoid flattening, are used extensively in the computations associated with the Loran navigation systems. Lambert never published his derivation and Andoyer's, published in Bulletin Géodésique 1932, does not indicate the extension to higher order terms in the flattening. Several new derivations are given with extension to higher order terms in the flattening, including identification of the forms in an 1895 paper of A. R. Forsyth.

6. *The application of functional analysis to the solution of differential and integral equations*, by Seymour Goldberg, University of Maryland, College Park, Maryland, (invited address).

A survey is given concerning the progress of the theory of unbounded linear operators in Banach spaces. The perturbation theorems of Kato and Gokhberg-Krein are presented with applications to differential and integral equations.

7. *The converse of Banach's contraction theorem*, by Philip R. Meyers, National Bureau of Standards, Washington, D. C.

After adding one simple conclusion of Banach's contraction theorem to the usual ones, the speaker shows that "f is a contraction for some complete metric on the topological space X" is a necessary condition for these conclusions. This constitutes a converse of the contraction theorem.

8. *Some observations on mathematics in the Middle East*, by Daniel B. Lloyd, District of Columbia Teachers College, Washington, D. C.

Recent archaeological discoveries in Iraq reveal a depth of understanding by the Babylonians

of (circa) 1700 B.C. concerning principles of Euclidean geometry formerly attributed solely to the Greeks of a later era. Two of the latest mathematical tablets unearthed contain unique and unusual solutions of triangle areas and their relationships.

9. *A variant of the two-dimensional Riemann integral,* by A. J. Goldman, National Bureau of Standards, Washington, D. C.

S. Marcus (this MONTHLY, 71 (1964) 544–545) proposed a variant of the two-dimensional Riemann integral and observed that $f(x, y) = x$ failed to be integrable. Here a necessary and sufficient condition is given for integrability, the gist of which is that the function must be very nearly constant. In particular, the constants are the only continuous integrable functions.

10. *Commutative semi-groups and pseudo-inverses,* by Kenneth D. Taylor, Army Map Service, Washington, D. C.

If c is such that $acb = ab$ for all a, b in the semi-group, then c is called a pseudo-inverse. The theorems and corollaries center about the principal theorem: If G is a semi-group with an element d in G such that for all a, b in G $adb = ba$, then the semi-group is commutative. The maximal sub-semi-group of G which is a group, is found. Finally, the question of whether an extension will produce a pseudo-inverse is discussed.

S. S. SASLAW, *Secretary*

MAY MEETING OF THE MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA SECTION

The annual Spring meeting of the Maryland-District of Columbia-Virginia Section of the MAA was held at the University of Maryland, College Park, Maryland, on May 1, 1965. Dr. Daniel Shanks, Chairman of the section presided. There were 110 persons in attendance, including 96 members of the Association. Very lively discussions followed both the morning and afternoon sessions.

At the business meeting the following officers were elected:

Chairman, Professor Samuel S. Saslaw, United States Naval Academy, Annapolis, Maryland; Vice Chairmen, Professor Hyman Kamel, Howard University, Washington, D. C.; and Professor Thomas L. Reynolds, College of William and Mary, Williamsburg, Virginia; Secretary, Professor George N. Trytten, University of Maryland, College Park, Maryland; Treasurer, Professor Stanley B. Jackson, University of Maryland, College Park, Maryland.

The following program was presented:

1. *The aims and purposes of the Committee on the Undergraduate Program in Mathematics of the Mathematical Association of America,* by Samuel Eilenberg, Columbia University (invited address).

2. *Question and discussion period.*

3. *What Graduate Schools in the Area Expect of New Students,* Panel discussion. Participants: Gustav B. Hensel, Catholic University of America; John D. Miller, University of Virginia; Bruce Reinhart, University of Maryland.

4. *Question and Discussion period.*

S. S. SASLAW, *Secretary*

NOVEMBER MEETING OF THE INDIANA SECTION

The Indiana Section of the MAA met on Saturday, November 6, 1965, at Franklin College, Franklin. Approximately 100 persons attended, of whom 67 were members of the Association. Chairman George Springer of Indiana University presided. The morning was devoted to short papers and the afternoon to a business meeting and an invited hour address entitled "Categories" by Professor Samuel Eilenberg of Columbia University.

Papers presented at the morning session were:

1. *Fiducial theory and invariant estimation*, by R. B. Hora, Indianapolis Branch, Purdue University, and R. J. Buehler, University of Minnesota. (Report by R. B. Hora.)

Let P^{ω} be a family of distributions which satisfy assumptions that are essentially equivalent to those of Fraser (Biometrika, 48 (1962) 261) but are closer to those of Stone (Ann. Math. Stat., 36 (1965) 440). Let E_f denote fiducial expectation and E_R the conditional expectation given the ancillary. It is shown that for invariant functions $H(\chi, \omega)$, the identity $E_fH = E_RH$ holds. The identity has been applied to obtain best invariant estimators for "invariantly estimable functions." Finally, the relations between estimation of "invariantly estimable functions" and coset estimation and fiducial and confidence limits have been considered.

2. *Programmed instruction in college level mathematics*, by Thomas A. Davis, DePauw University.

This was a discussion of the use of programmed instruction in pre-calculus mathematics at DePauw University during 1964–65. The program included such topics as inequalities, absolute values, symmetry, asymptotes, and equations of straight lines. The performance of students who used the programmed material was slightly better than that of students who learned the same material by the traditional classroom lecture and discussion method. The time spent on the program was about one-half the time spent by the students who learned the material in the classroom.

3. *Binary relations as the basis for a fact retrieval system*, by Roger Elliott, Indiana State University.

A model for a computer fact retrieval system is developed. The model accepts facts and queries couched in a relational language, and it constructs data structures which are efficient for storing the input data and which permit efficient inference-making. Data are presented verifying an hypothesized independence of retrieval time of the size of the data base.

4. *Extremal properties of spline interpolants*, by Michael Golomb, Purdue University.

An extremal property possessed by interpolating spline functions was derived which includes all the minimizing properties of these approximants found previously (Schoenberg, Sard, Golomb-Weinberger, Holladay, et al.).

5. *On regular rings*, by Jiang Luh, Indiana State University. Following O. Steinfeld, a subring Q of a ring A is said to be a quasi-ideal of A if $AQ \cap QA \subseteq Q$. In this paper characterizations for regular rings and for strongly regular rings are given in terms of quasi-ideals.

6. *Behavior of derivatives*, by A. Bruckner, University of California at Santa Barbara, visiting Professor at Purdue University.

The speaker gave an exposition of classical examples of functions possessing various properties with respect to differentiability.

7. *On the Hodgkin-Huxley partial differential equation*, by H. Melvin Lieberstein, Indiana University.

The equations for propagation of impulses on an unmyelinated squid axon are modified to include effects of core capacitance and inductance without introducing any new parameters. One equation is the one dimensional wave operator acting on a function whose values represent voltage across a surrounding membrane, set equal to a non-linear function of membrane voltage and its first time derivative. This is taken together with the first order differential equations for the Hodgkin-Huxley empirical parameters n, m and h. Voltage is specified as a $-15mv$ sawtooth or square impulse at a cut end and is required to be initially zero together with its first time derivative along the axon. This generates numerically a -100 or a $-103mv$ action potential. Initial waves of $-10mv$ die out. A wave which propagates with velocity and form constant to five digits develops as x and t increase; it satisfies a first order ordinary differential equation which replaces the former second order equation. Agreement with former calculations is excellent and the extreme sensitivity

to propagation rate is removed. The boundary value problem can be regarded as a model for transient development and propagation on a human muscle fiber. (The above work has been supported by NIH grant HE 10034).

Local arrangements for the meeting were in charge of Rodney Hood of Franklin College. President Wesley N. Haines of Franklin College welcomed participants on behalf of the college.

Six charter members of the Association now residing in Indiana were accorded special recognition at the meeting. These men are Will E. Edington, Gordon H. Graves, Paul R. Rider, Charles K. Robbins, Clarence P. Sousley and Harold E. Wolfe. Of these only Messrs. Graves and Rider were able to attend the meeting.

P. T. MIELKE, *Secretary*

NOVEMBER MEETING OF THE NEW JERSEY SECTION

The tenth annual meeting of the New Jersey Section of the MAA was held at Montclair State College, Upper Montclair, on November 6, 1965. Professor Max A. Sobel, Chairman of the Section, presided at the morning session. Professor Joshua Barlaz presided at the afternoon session. One hundred and forty-three persons attended the meeting, including one hundred and twelve members of the Association.

At the business meeting the following members were elected: Professor Joshua Barlaz of Rutgers University, Chairman of the Section (Nov. '66); Professor Bernard Greenspan of Drew University, Member-at-Large of the Executive Committee (Nov. '68); Professor Hale Trotter of Princeton University, Member-at-Large of the Executive Committee (Nov. '66); Professor John K. Reckzeh of Jersey City State College, Associate Secretary-Treasurer (Nov. '68); Professor Francis A. Varrichio of Saint Peter's College, Secretary-Treasurer. Reports were given by L. F. McAuley, Governor of the Section; J. Barlaz, Section Representative to the Summer meeting; W. A. Krzeminski, member of the High School Contest Committee; F. A. Varrichio, Secretary-Treasurer.

At the morning session the following papers were presented:

1. *Logic: Fad or Tool?* by Prof. Hassler Whitney, Institute of Advanced Study, Princeton, N. J. (By invitation.)

Many texts in general mathematics contain a chapter on logic; it commonly seems unconnected with actual mathematics, and is likely not to be made use of later. Yet all mathematics uses logical reasoning throughout. What is really needed is to understand the elementary principles used in actually carrying on mathematics. This involves not only propositions and quantifiers, but also an understanding of the use of symbols, the meaning of statements relative to the underlying hypotheses, etc. Various simple examples are given to illustrate these ideas.

2. *Countable Topological Spaces*, by Albert Wilansky of Lehigh University, Bethlehem, Pa. (By invitation.)

Let X be a countable T_1 space. If regular, it is normal, zero-dimensional, totally disconnected. If regular and first countable or locally compact, it is metrizable, indeed a subset of the rationals. Let $\beta = N \cup \{t\}$, $t \epsilon \beta N/N$. Then β is normal, not first countable, pseudofinite, hemicompact. Any hemicompact first countable space is locally compact. Let $X = \{t\} \cup N$ with N discrete, $t \epsilon \overline{N}$; deleted neighborhoods of 0 being $\{S: \lim_A \chi(S) = 1\}$, A a fixed positive regular matrix. A. K. Snyder has associated topological properties of X with summability properties of A.

The afternoon session consisted of a panel discussion of *The Twelfth Year Program in High School Mathematics*. The panelists were: Mr. Martin Moskowitz of Vailsburg High School, Newark, N. J., Mr. Henry Peterson, Wayne Senior High School, Wayne, N. J., and Dr. Anthony Pettofrezzo, Montclair State College.

F. A. VARRICHIO, *Secretary*

CALENDAR OF FUTURE MEETINGS

Forty-seventh Summer Meeting, Rutgers, The State University, New Brunswick, New Jersey, August 29–31, 1966.

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN, Waynesburg College, Waynesburg, Pennsylvania, April 30, 1966.

ILLINOIS, Saint Dominic College, St. Charles, May 13–14, 1966.

INDIANA, Indiana State University, Terre Haute, May 14, 1966.

IOWA, Central College, Pella, April 15, 1966.

KANSAS, University of Kansas, Lawrence, March 26, 1966.

KENTUCKY, University of Kentucky, Lexington, April 29–30, 1966.

LOUISIANA-MISSISSIPPI

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, College of William and Mary, Williamsburg, Va., April 30, 1966.

METROPOLITAN NEW YORK, St. John's University, Jamaica, N. Y., May 7, 1966.

MICHIGAN, Wayne State University, Detroit, Michigan, April 2, 1966.

MINNESOTA, Macalester College, St. Paul, May 1966.

MISSOURI, University of Missouri at Rolla, April 30, 1966.

NEBRASKA, Nebraska Center for Continuing Education, Lincoln, April 29–30, 1966.

NEW JERSEY, Rutgers, The State University, New Brunswick, November 12, 1966.

NORTHEASTERN

NORTHERN CALIFORNIA

OHIO, Ohio Wesleyan University, Delaware, April 23, 1966.

OKLAHOMA-ARKANSAS, Oklahoma Baptist University, Shawnee, April 1–2, 1966.

PACIFIC NORTHWEST, University of Victoria, Victoria, British Columbia, June 17, 1966.

PHILADELPHIA, Villanova University, Villanova, November, 1966.

ROCKY MOUNTAIN, Colorado State University, Fort Collins, May 13–14, 1966.

SOUTHEASTERN, Emory University, Atlanta, Georgia, March 25–26, 1966.

SOUTHERN CALIFORNIA, Occidental College, Los Angeles, March 12, 1966.

SOUTHWESTERN, University of New Mexico, Albuquerque, April 1–2, 1966.

TEXAS, Southern Methodist University, Dallas, April 15–16, 1966.

UPPER NEW YORK STATE, St. Bonaventure University, Olean, May 14, 1966.

WISCONSIN, Wisconsin State University, Eau Claire, May 7, 1966.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, New Brunswick, N. J., August 30–September 2, 1966.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Washington State University, Pullman, June 20–24, 1966.

ASSOCIATION FOR COMPUTING MACHINERY, Ambassador Hotel, Los Angeles, August 30-September 1, 1966.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

INSTITUTE OF MATHEMATICAL STATISTICS, Rutgers, The State University, New Brunswick, N. J., August 30–September 2, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Americana Hotel, New York City, April 13–16, 1966.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Miramar Hotel, Santa Monica, California, May 18–20, 1966.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Iowa, Iowa City, May 11–14, 1966. (Symposium on Numerical Analysis.)

THE AMERICAN
MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 3

CONTENTS

MARCH 1966

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

THIS IS THE OFFICIAL JOURNAL OF THE

MATHEMATICAL ASSOCIATION OF AMERICA, INC.

(Devoted to the Interests of Collegiate Mathematics)

General Office: SUNY at Buffalo, Buffalo, N. Y. 14214. *Executive Director:* H. M. GEHMAN

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

THE THEORY OF ROUND ROBIN TOURNAMENTS

FRANK HARARY, University of Michigan, and
LEO MOSER, University of Alberta

In this review paper, we make a detailed study of a class of directed graphs, known as tournaments. The reason they are called tournaments is that they represent the structure of round robin tournaments, in which players or teams engage in a game that cannot end in a tie and in which every player plays each other exactly once.

Although tournaments are quite restricted structurally, they are realized by a great many empirical phenomena in addition to round robin competitions. For example, it is known that many species of birds and mammals develop dominance relations so that for every pair of individuals, one dominates the other. Thus, the digraph of the "pecking structure" of a flock of hens is asymmetric and complete, and hence a tournament.

Still another realization of tournaments arises in the method of scaling, known as "paired comparisons." Suppose, for example, that one wants to know the structure of a person's preferences among a collection of competing brands of a product. He can be asked to indicate for each pair of brands which one he prefers. If he is not allowed to indicate indifference, the structure of his stated preferences can be represented by a tournament.

Tournaments appear similarly in the theory of committees and elections. Suppose that a committee is considering four alternative policies. It has been argued that the best decision will be reached by a series of votes in which each policy is paired against each other. The outcome of these votes can be represented by a digraph whose points are policies and whose lines indicate that one policy defeated the other. Such a digraph is clearly a tournament.

After giving some essential definitions, we develop properties that all tournaments display. We then turn our attention to transitive tournaments, namely those that are complete orders. It is well known that not all preference structures are transitive. There is considerable interest, therefore, in knowing how transitive any given tournament is. Such an index is presented toward the end of the second section. In the final section, we consider some properties of strongly connected tournaments.

1. Definitions and preliminary concepts. It is necessary, alas, to include quite a few definitions, to make the treatment precise and self-contained.

A *directed graph*, or *digraph* D for short, consists of a finite set $V=\{v_1, v_2, \cdots, v_p\}$ of *points* together with a subset of $V\times V$, whose elements are called *lines*. Each line $(v_i, v_j)=v_iv_j$ is directed and goes from its first point v_i to a different second point v_j, so that a digraph is *irreflexive*. A digraph is *asymmetric* if whenever line v_iv_j is in it, then v_jv_i is not. It is *transitive* if for every three distinct points v_i, v_j, v_k, the existence of lines v_iv_j and v_jv_k implies the existence

of line v_iv_k. In a *complete* digraph, for any two distinct points v_i and v_j, line v_iv_j or line v_jv_i exists.

A *subgraph* of a digraph D is a subset of the points and lines of D which themselves form a digraph. The *removal of a point* v from D results in the maximal subgraph, D-v, not containing v. That is, D-v has all points of D except v and all lines except those to and from v. If U is a proper subset of the set V of points of D, then D-U is the digraph obtained by removing the points of U in succession. The *subgraph* $\langle U \rangle$ *generated by a set* U *of points* of D contains the points of U and those lines of D from one point of U to another.

If uv is a line of a digraph, then u is said to be *adjacent to* v and v is *adjacent from* u. The *outdegree*, denoted od v, of a point v is the number of points adjacent from v; its *indegree*, id v, is the number of points adjacent to v. A *transmitter* is a point with positive outdegree and zero indegree; *a receiver* has positive indegree and zero outdegree.

A *walk from* u *to* v is an alternating sequence of points and lines of the form $u_1, u_1u_2, u_2, u_2u_3, \cdots, u_{n-1}u_n, u_n$, in which $u=u_1$ and $v=u_n$. For brevity this is written $u_1u_2 \cdots u_n$ since then the lines are clear from context. A walk in which all points (and hence all lines) are distinct is called a *path*. A walk of positive length in which only the first and last points are the same is called a *cycle*. The *length* of a path or a cycle is the number of lines in it. A *complete path* or *cycle* contains all the points of the given digraph. If there is a path from u to v, then v is *reachable from* u. The *distance* from u to v, denoted $d(u, v)$, is the length of a shortest such path.

A digraph is called *strongly connected*, or *strong*, if every pair of points are mutually reachable. A *strong component* of a digraph is a maximal strongly connected subgraph. The *condensation* D^* of digraph D has as its points the strong components of D; there is a line in D^* from one strong component S_i to another S_j if there is a line in D from some point of S_i to a point of S_j.

The *converse* D' of a digraph D has the same points as D and contains line v_iv_j if and only if v_jv_i is a line of D. In other words, D' is obtained from D by reversing the direction of every line. Every concept in directed graph theory has a converse concept. For example, the outdegree and indegree of a point are converse concepts of each other, as are "adjacent to" and "adjacent from," etc. A point and a line are their own converses. A valuable principle in the theory of directed graphs is the following, borrowed from the theory of relations.

Directional Duality Principle. For each theorem about digraphs, there is a corresponding theorem which is obtained by replacing every concept by its converse.

2. Some properties of tournaments. A *tournament* is a complete asymmetric digraph. A brief discussion of tournaments is given by Ore in [16]. The smallest tournaments are shown in Figure 1. Clearly, if U is a proper subset of the points in a tournament T, then T-U is also a tournament.

It follows directly from the definitions that if v is any point in a tournament

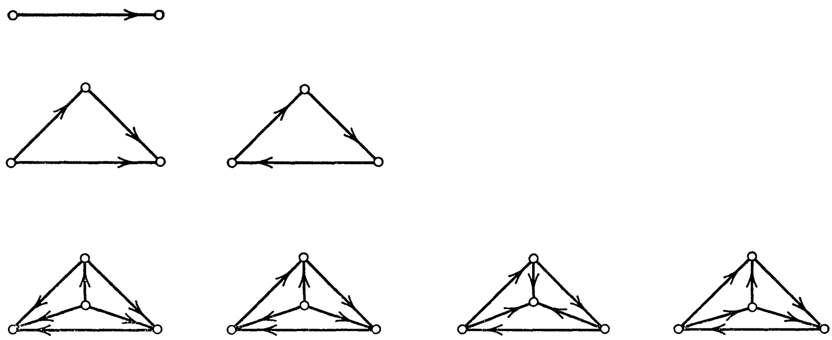


FIG. 1

T with p points, then od $v+$id $v=p-1$. Also, the total number of lines in T is $\frac{1}{2}p(p-1)$. In the tournament of a round robin competition, the outdegree of a point is the number of victories won by that player. For this reason, we shall call the outdegree of a point v_i of a tournament its *score*, denoted s_i.

The *score sequence* of a tournament T is the ordered sequence of integers $(s_1, s_2, \cdots, s_p)$. We assume without loss of generality that the points v_i have been ordered in such a way that $s_1 \leqq s_2 \leqq \cdots \leqq s_p$. The following theorem by Landau [13] gives a necessary and sufficient condition for a sequence of nonnegative integers to be the scores of some tournament.

THEOREM 1. *A sequence of nonnegative integers $s_1 \leqq s_2 \leqq \cdots \leqq s_p$ is a score sequence if and only if their sum satisfies the equation:*

$$\text{(I)} \qquad \sum_{i=1}^{p} s_i = \tfrac{1}{2}p(p-1),$$

and the following inequalities hold for every positive integer $k<p$:

$$\text{(II)} \qquad \sum_{i=1}^{k} s_i \geqq \tfrac{1}{2}k(k-1).$$

We prove the necessity of conditions (I) and (II) by taking $s_1 \leqq s_2 \leqq \cdots \leqq s_p$ as the score sequence of a tournament T. Since the sum of the scores of T is the number q of lines, and since $q=\frac{1}{2}p(p-1)$, equation (I) is verified. To establish the inequalities (II), we note that, for any integer $k<p$, the subtournament of T whose points are $v_1, v_2, \cdots, v_k$ contains exactly $\frac{1}{2}k(k-1)$ lines. Hence in the entire tournament T, $\sum_{i=1}^{k} s_i \geqq \frac{1}{2}k(k-1)$, since there may occur in T a line to one of the other $p-k$ points. The proof of the converse is considerably more involved, and is omitted.

Moon has proved in [15] that Theorem 1 can be generalized to tournaments with non-integral scores, with conditions (I) and (II) still serving as a criterion for a score sequence.

Theorem 1 may be illustrated by the tournament with five points shown in Figure 2. Clearly, the score sequence of this tournament is (1, 1, 2, 3, 3,). It is immediately apparent that equation (I) and the inequalities of (II) are satisfied.

Consider a basketball league consisting of ten teams in which each team plays every other team once. Since no game can end in a tie, the digraph of the outcomes of all games at the end of the season is a tournament. What are the possible distributions of the number of victories among the teams? Clearly, each distribution must satisfy conditions (I) and (II) of Theorem 1. This fact provides information concerning certain questions that may be asked about the final standings of the team. For example, what is the largest number of teams that can have a winning season? The answer is nine, since the sequence of integers (0, 5, 5, 5, 5, 5, 5, 5, 5, 5) satisfies the conditions of Theorem 1 and no sequence containing ten integers, all greater than 4, does. Can the season end in a complete tie? Clearly not, since the average of the scores is not an integer.

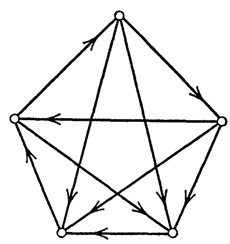


FIG. 2

If a tournament T has a transmitter v, then v is adjacent to every point of T. Thus, its distance to every point is 1, and its score is $p-1$. Clearly a tournament has at most one transmitter. If T does not have a transmitter, there will, of course, be at least one point with maximum score, which is less than $p-1$. The next theorem provides information concerning the location of such points in T; it is a "folk-theorem" cited in [10].

THEOREM 2. *In any tournament, the distance from a point with maximum score to any other point is* 1 *or* 2.

To prove this theorem, let v be any point whose score s is maximum. Without loss of generality, we denote the points to which v is adjacent by $v_1, v_2, \cdots, v_s$. Since T is a tournament, v is adjacent from the remaining $p-1-s$ points $u_1, u_2, \cdots, u_{p-1-s}$, as in Figure 3. The proof will be complete if we show that each point u_k is adjacent from at least one point v_j, for then each distance $d(v, v_j)=1$ and $d(v, u_k)=2$. Assume that this is not the case for the point u_1. Then u_1 is adjacent to every point $v_1, v_2, \cdots, v_s$ as well as to v itself. Hence, its score is $\mathrm{od}(u_1) \geqq s+1$. This contradicts the hypothesis that s is the largest score of any point.

An interesting consequence of Theorem 2 follows when it is applied to a round robin competition. Let v be a player with maximum score in such a com-

petition consisting of at least three players. Then any player who defeats v is himself defeated by another player defeated by v. The directional dual of this conclusion says that if v is a player with minimum score, then any player defeated by v defeats another player who defeats v.

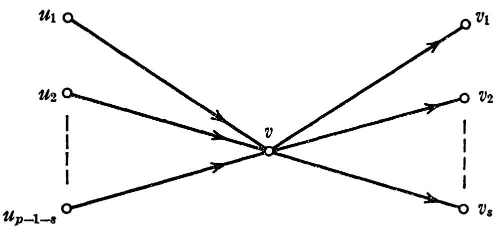


FIG. 3

The directional dual of the next theorem was stated by Silverman [**18**] in the following picturesque terminology: Consider a club in which among any two members, one is a creditor of the other. A "bum" is defined as a member who is in debt to everyone else. A "deadbeat" is not a "bum," but for each member he does not owe, he owes someone who owes this member. Then if the club has no bums, it has at least three deadbeats. In the future, the directional duals of theorems will not be given, but the reader is urged to develop them for himself.

THEOREM 3. *If a tournament has no transmitter, then it contains at least three points each of which can reach every point in at most two steps.*

Let u be a point of T with maximum score. By Theorem 2, u can reach every point in at most 2 steps. Since T has no transmitter, there is at least one point adjacent to u. Among all such points, let v have maximum score. Suppose there is a point v_0 not reachable from v within 2 steps. It follows that v_0 is necessarily adjacent to v and to every point which is adjacent from v, in particular, the point u. But then $s(v_0) \geqq s(v)+1$, contradicting the choice of v. Therefore every point is reachable from v in at most 2 steps. By the same argument, if w has greatest score among the points adjacent to v, then w can reach every point with 2 steps. Since T is asymmetric and lines wv and vu are in T, necessarily u, v, and w are distinct points. Since every point is within distance 2 from each of these, the theorem is proved.

The next theorem due to Rédei [**17**] is certainly the best known result concerning tournaments. It also holds for any complete digraph, as stated by König in [**12**]. Actually, Rédei showed that every tournament has an odd number of complete paths.

THEOREM 4. *Every tournament has a complete path.*

The proof is given by induction on the number p of points. Referring to Figure 2, we see that every tournament with 2, 3, or 4 points has a complete

path. As the inductive hypothesis, let the theorem hold for all tournaments with n points. Let T be any tournament with $n+1$ points. To complete the proof of the theorem, it is necessary to show that T has a complete path.

Let v_0 be any point of T. Then $T-v_0$ is a tournament with n points. Since the inductive hypothesis applies to $T-v_0$, it has a complete path which may be denoted by $P=v_1v_2v_3 \cdots v_n$. Let us return to T and see how the point v_0 can be added to P in order to obtain a complete path of T. Consider the two points v_0 and v_1 of T. There are two possibilities: either line v_0v_1 or v_1v_0 is in T. If v_0v_1 is a line of T, then $v_0v_1v_2v_3 \cdots v_n$ is a complete path of T. On the other hand, if v_1v_0 is in T, then let v_i be the first point of P, if any, for which the line v_0v_i is in T. Then necessarily line $v_{i-1}v_0$ is in T. Therefore $v_1v_2 \cdots v_{i-1}v_0v_i \cdots v_n$ is a complete path of T (as shown in Figure 4). But there may not be any such first point v_i, since v_0 might be a receiver of T. In that case, $v_1v_2v_3 \cdots v_nv_0$ is a complete path of T, completing the proof.

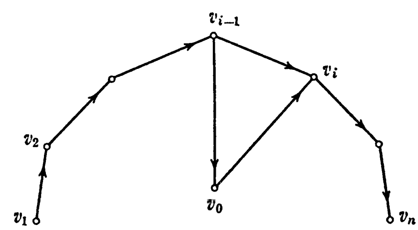


FIG. 4

Since every tournament has a complete path, it is possible to order all the players in a round robin competition so that each defeats the succeeding one. Thus the integers $1, 2, \cdots, p$ can be assigned to the players to indicate their rank in this order. There are, however, two serious difficulties in such a procedure. First, there is no necessary relation, in general, between such a ranking of players and their scores. Second, a tournament may have more than one complete path, so that several different rankings may be possible. Figure 2 illustrates these observations. For example, there is a complete path from one of the two points with lowest score to the other one. In fact, each point has every possible rank in some complete path of this tournament.

Another way of stating these difficulties is to say that a tournament need not be a complete order. For if there is a complete order on a set of p points, then there exists a one-to-one correspondence between these points and the integers $1, 2, \cdots, p$, in their natural order. Thus, whenever there is a complete order on a set of points, each point can be assigned a distinct rank. A complete order is a relation: irreflexive, asymmetric, complete, and transitive. Since every tournament has the first three of these properties, a transitive tournament is a complete order. There is considerable interest, therefore, in knowing the properties of transitive tournaments and, for any particular tournament, how much transitivity it displays.

How Transitive is a Tournament?

In analyzing the degree of transitivity of a tournament, it is useful to refer to the subtournament generated by any three of its points. A *triple* $\langle uvw \rangle$ of a tournament is the subtournament generated by the points u, v, and w. We saw in Figure 1 that there are only two tournaments with three points; one of these is transitive and the other is cyclic. Obviously every triple in a transitive tournament is transitive. Also, if a tournament is not transitive, it must contain at least one cyclic triple. It is possible therefore to quantify its degree of transitivity by using the number of triples of each kind.

Before dealing directly with the question of the degree of transitivity of a tournament, we first state the well-known structural characterization of a complete order. If v is any point of a complete order with at least three points, then $T-v$ is also a complete order. Also, every complete order has a unique transmitter and a unique receiver.

THEOREM 5. *If T is a complete order with p points, then T is isomorphic with the tournament T_p whose points are $v_1, v_2, \cdots, v_p$ in which v_i is adjacent to v_j if and only if $i<j$.*

The first corollary of this theorem lists without proof the major equivalent characterizations of a complete order.

COROLLARY 5a. *The following statements are equivalent for any tournament T with p points.*

(1) *T is transitive.*
(2) *T is acyclic.*
(3) *T has a unique complete path.*
(4) *The score sequence of T is $(0, 1, 2, \cdots, p-1)$.*
(5) *T has $p(p-1)(p-2)/6$ transitive triples.*

COROLLARY 5b. *The condensation of a tournament is a complete order.*

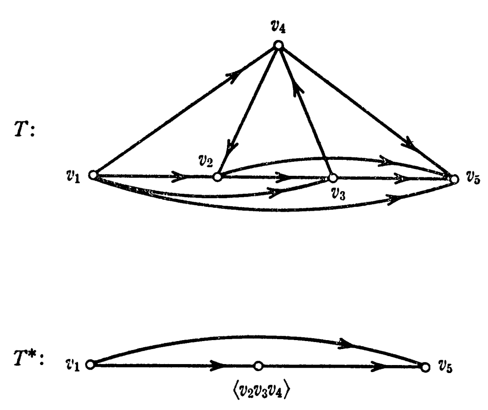


FIG. 5

If T is strong, then T^* is a single point. Therefore, take T as not strong. Obviously the condensation of a tournament is itself a tournament. And since T^* is acyclic, it is transitive.

Figure 5 shows a tournament T and its condensation T^*. Clearly, T is not transitive but T^* is.

If a tournament is not transitive, then it is often useful to know how many transitive triples it has. The next theorem shows that the number of transitive triples in any tournament may be easily calculated from the score sequence of the tournament. Theorem 6 and its first two corollaries are implicitly contained in Kendall and Babington Smith [**11**]; Corollary 6c appears there explicitly.

THEOREM 6. *The number b of transitive triples in a tournament T with score sequence $(s_1, s_2, \cdots, s_p)$ is*

$$b = \sum_{i=1}^{p} \tfrac{1}{2}s_i(s_i - 1).$$

To prove this formula, let $b_i = \frac{1}{2}s_i(s_i - 1)$. Then b_i is the number of combinations of s_i objects taken two at a time. But s_i is the number of points adjacent from v_i. Therefore b_i is the number of pairs of points adjacent from v_i. But any transitive triple in the tournament T has a unique transmitter within it. Hence b_i is the number of transitive triples whose transmitter is v_i. Clearly, the number b of transitive triples in T is obtained by adding these number b_i for all points, proving the theorem.

A little algebraic manipulation transforms the equation of Theorem 6 into the following equivalent form.

COROLLARY 6a. *The number b of transitive triples in T is*

$$b = \tfrac{1}{2} \sum_{i=1}^{p} s_i^2 - \tfrac{1}{4}\, p(p - 1).$$

Since the total number of triples in any tournament with p points is $\binom{p}{3}$, and since each triple is either transitive or cyclic, we obtain the following formula for the number of cyclic triples.

COROLLARY 6b. *The number c of cyclic triples in a tournament satisfies the equation*

$$c = \tfrac{1}{6}\, p(p - 1)(p - 2) - \tfrac{1}{2} \sum_{i=1}^{p} s_i(s_i - 1).$$

The next corollary gives the maximum number of cyclic triples that can occur in any tournament with a given number of points. Its proof is omitted.

COROLLARY 6c. *Among all the tournaments with p points, the maximum number of cyclic triples is*

$$c_{\max}(p) = \begin{cases} \dfrac{p^3 - p}{24} & \textit{if } p \textit{ is odd, and} \\ \dfrac{p^3 - 4p}{24} & \textit{if } p \textit{ is even.} \end{cases}$$

Very recently, we (Beineke and Harary [2]) have generalized the result of the preceding corollary by obtaining a formula for the maximum number of subsets of n points which generate a strongly connected subtournament that can occur in any tournament with a given number p of points. Using the methods of flows in networks as developed in Ford and Fulkerson [6], a recent result of Fulkerson [8] determines the maximum possible number of upsets (defeat of a player by another player of lower score) that can occur in a round robin tournament with a given score sequence.

Coefficient of Consistency. In research employing paired comparisons, it is usually assumed that if a judge is entirely "consistent" in his decisions, the result will be a complete order, which therefore has no cyclic triples. Since in actual practice judges are seldom completely consistent, it is useful to have a coefficient indicative of the degree of consistency among the comparisons. Such a coefficient has been proposed by Kendall and Babington Smith [11]. They wanted their coefficient of consistency to be normalized in such a way that its value is 1 when the tournament of comparisons is transitive and 0 when the tournament contains as many cyclic triples as possible, i.e., when it is as inconsistent as possible. Making use of the results in Corollary 6c, they define a coefficient of consistency ξ by the following equation in which c is the number of cyclic triples of a given tournament T with p points resulting from paired comparisons,

$$\xi = 1 - \frac{c}{c_{\max}(p)} \cdot$$

Thus we see that $0 \leqq \xi \leqq 1$ for any tournament T. On substituting the result of Corollary 6c, this equation becomes:

$$\xi = \begin{cases} 1 - \dfrac{24c}{p^3 - p} & \text{when } p \text{ is odd, and} \\ 1 - \dfrac{24c}{p^3 - 4p} & \text{when } p \text{ is even.} \end{cases}$$

Another coefficient of consistency has been proposed by Berge [3]. His ratio for measuring consistency is $b/\binom{p}{3}$, where b is the number of transitive triples of a tournament T, and the denominator $\binom{p}{3}$ is the total number of triples. This ratio will take on the value 1 if and only if T is transitive. But it will never have the value 0 when $p > 3$ since every tournament with more than 3 points must contain at least one transitive triple.

Unsolved Problem. To each tournament T, one can assign a positive integer trans (T) which gives the order of a largest transitive subtournament contained in T. What is the minimum value $f(p)$ of trans (T) as T varies over all tournaments with p points, expressed as an explicit function of p? In an unpublished work, P. Erdös and L. Moser have shown that this value has the order of magnitude $c \log_2 p$ for large p, but the constant c is undetermined. From Figure 1, we see that $f(2) = f(3) = 2$ and $f(4) = 3$.

The Voting Paradox. Consider an electorate which, by majority vote, is to choose among a set of motions or candidates. Assume that the chairman casts a deciding vote in case of a tie. The possible outcomes resulting from the pairing of each motion against each other can be represented by a tournament in which each motion corresponds to a point and the fact that motion v_i can defeat motion v_j corresponds to a line v_iv_j. It has long been known that such a tournament may contain a cycle. In fact, the term "cyclical majorities" was employed by the Rev. C. L. Dodgson (Lewis Carroll), cf. Black [4], to describe this kind of situation.

This voting paradox has stimulated a considerable literature concerning which procedure is "best," but no completely satisfactory method has been devised. Arrow [1] has considered the more general problem of finding a "social welfare function," whereby the preferences of individuals are equitably combined into a preference ordering by society. He has shown that it is impossible to satisfy one set of five plausible conditions for an equitable welfare function. A discussion of Arrow's Impossibility Theorem may be found in Luce and Raiffa [14].

Strong Tournaments

In Corollary 5b, we saw that the condensation of any tournament is a complete order. Thus if a tournament is not transitive, one may wish to study its strong components. For this reason, there is considerable interest in investigating the properties of strong tournaments. In this section we present two criteria for a tournament to be strong as well as several conditions which are sufficient but not necessary.

THEOREM 7. *If a tournament T is strong, then it contains a cycle of each length* $k = 3, 4, \cdots, p$.

This theorem is proved by induction. First we observe that since T is not transitive, it has a cyclic triple. We now take as our inductive hypothesis the statement that T has a cycle Z of length $k < p$. We will show that T must have a cycle of length $k+1$. Let us label the cycle Z as $v_1v_2v_3 \cdots v_kv_1$. There are just two possibilities. Either there is a point u not in Z such that u is adjacent to some point of Z and adjacent from another point of Z, or there is no such point u.

Case 1. There is a point u not in Z such that for some points v and w in Z, lines uv and wu occur in T. The proof for this case is illustrated in Figure 6. Let

us assume that the line v_1u is in T. Let v_i be the first point, going around the cycle from v_1, such that the line uv_i is in T. Then line $v_{i-1}u$ must be in T. Thus we see that T contains a cycle of length $k+1$, namely $v_1v_2v_3 \cdots v_{i-1}uv_i \cdots v_kv_1$.

Case 2. There is no point u as in Case 1. In this case, all points of T which are not in Z can be partitioned into two subsets U_1 and U_2, where U_1 is the set of all points not in Z adjacent to (every point of) Z, and U_2 is the set of all points not in Z adjacent from (every point of) Z. The sets U_1 and U_2 are not empty since T is strong by hypothesis, and they are disjoint by the hypothesis of Case 2. Since T is strong, there exist points u_1 in U_1 and u_2 in U_2 such that line u_2u_1 is in T. Then we may again construct a cycle of length $k+1$ in T, namely: $u_1v_1v_2 \cdots v_{k-1}u_2u_1$.

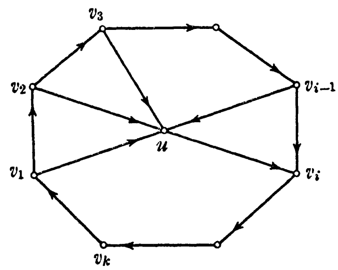


FIG. 6

By proving for each of these two cases that if T has a cycle of length k, it has a cycle of length $k+1$, we have completed the proof of the theorem.

By Corollary 5a, one criterion for a tournament to be transitive is that it have a unique complete path. The first corollary of Theorem 7 provides an analogous criterion for a tournament to be strong, cf. Camion [5] and Foulkes [7].

COROLLARY 7a. *A tournament is strong if and only if it has a complete cycle.*

By Theorem 7, every strong tournament with p points has a cycle of length p, which is a complete cycle. If a digraph has a complete cycle, it is obviously strong.

COROLLARY 7b. *If T is a strong tournament, it has at least $p-2$ cyclic triples.*

We prove the result by induction. If a strong tournament has exactly 3 points, it clearly has $p-2$ cyclic triples. As the inductive hypothesis, we take the result as true when $p=n$. Let T be strong with $n+1$ points. By Theorem 7, T has a cycle of length $p-1$. Hence, there is a point v_0 such that $T-v_0$ has a complete cycle $Z=v_1v_2 \cdots v_nv_1$. Because T is strong, v_0 has positive score. Without loss of generality we take the line v_0v_1 in T. Since v_0 also has positive indegree, we let v_i be the first point, going around the cycle Z from v_1, such that the line v_iv_0 is in T. Then $v_iv_0v_{i-1}v_i$ is a cycle so that T has at least one more cyclic triple than does $T-v_0$. But by the inductive hypothesis, $T-v_0$, having n

points and being strong, has at least $n-2$ cyclic triples. Hence T has at least $n-1$ cyclic triples, completing the proof.

COROLLARY 7c. *There exists a strong tournament of p points with $p-2$ cyclic triples.*

That no strong tournament can have fewer cyclic triples is shown in the preceding corollary. We now show that a tournament with this few cyclic triples can be constructed. Starting with a transitive tournament with p points, we replace the line from its transmitter u to its receiver v by vu forming a strong tournament T. Since $T-u$ and $T-v$ are transitive and therefore contain no cyclic triples, every cyclic triple of T contains both u and v and hence also vu. Every point of T other than u and v forms a cycle with vu. There are $p-2$ such points and hence $p-2$ cyclic triples in T.

The preceding corollaries give the number of cyclic triples necessary for a tournament to be strong. The next theorem, due to L. W. Beineke, gives a sufficient condition, in terms of the maximum number of cyclic triples in tournaments with p points, as in Corollary 6c.

THEOREM 8. *If T is a tournament with p points in which there are more cyclic triples than can occur in any tournament with $p-1$ points, then T is strong.*

Let T be a tournament with p points having more than $c_{\max}(p-1)$ cyclic triples. Suppose T is not strong, thus having at least two strong components. Clearly, all points of a cyclic triple lie in the same strong component. Hence if one strong component of T has k points, the number c of cyclic triples in T is no greater than the sum of the maximum numbers in tournaments with k and $p-k$ points: $c \leqq c_{\max}(k)+c_{\max}(p-k)$. However, it can be verified that if $0<k<p$, then $c_{\max}(p-1) \geqq c_{\max}(k)+c_{\max}(p-k)$. But this contradicts the assumption that $c>c_{\max}(p-1)$, thereby proving the theorem.

COROLLARY 8a. *If a tournament has p points and $c_{\max}(p)$ cyclic triples, then it is strong.*

It will be recalled that formulas (I) and (II) of Theorem 1 provide a necessary and sufficient condition for a sequence of nonnegative integers to be the scores of some tournament with p points. Formulas (I) and (II′) of the next theorem give the corresponding criterion for a strong tournament.

THEOREM 9. *Let T be a tournament with score sequence $s_1 \leqq s_2 \leqq \cdots \leqq s_p$. Then T is strong if and only if their sum satisfies the equation:*

$$\text{(I)} \qquad \sum_{i=1}^{p} s_i = \tfrac{1}{2}p(p-1),$$

and the following inequalities hold for every positive integer $k<p$:

$$\text{(II}'\text{)} \qquad \sum_{i=1}^{k} s_i > \tfrac{1}{2}k(k-1).$$

We first show that if T is a strong tournament, then conditions (I) and (II′) hold. We already know that condition (I) holds, since Theorem 1 has established it for any tournament. To verify the inequalities (II′), we note that for any integer $k<p$, the subtournament generated by $\{v_1, v_2, \cdots, v_k\}$ contains exactly $\frac{1}{2}k(k-1)$ lines. But since T is strong, there must be a line from one of these points to one of the $p-k$ points. Hence, in the entire tournament T, $\sum_1^k s_i > \frac{1}{2}k(k-1)$.

To prove the converse, consider conditions (I) and (II′) as given. We know by Theorem 1 that there exists a tournament T with these scores. Assume that such a tournament T is not strong. Then it has exactly one strong component S which is a receiver of the condensation T^*. Obviously the points in S have the smallest scores among all the points of T. If m is the number of points in S, then $m<p$ and $\sum_1^m s_i = \frac{1}{2}m(m-1)$, since there are no lines in T from a point in S to a point not in S. But one of the inequalities of the given condition (II′) is $\sum_1^m s_i > \frac{1}{2}m(m-1)$. This contradiction establishes the converse.

We next give some bounds for the scores of a tournament.

THEOREM 10. *Let T be a tournament with score sequence $s_1 \leqq s_2 \leqq \cdots \leqq s_p$. Then every score satisfies the inequalities: $\frac{1}{2}(k-1) \leqq s_k \leqq \frac{1}{2}(p+k-2)$.*

First, we suppose that $s_k < \frac{1}{2}(k-1)$. Then, for every $i<k$, $s_i \leqq s_k < \frac{1}{2}(k-1)$, so that $\sum_1^k s_i < \frac{1}{2}k(k-1)$. But by Theorem 1, $\sum_1^k s_i \geqq \frac{1}{2}k(k-1)$, which is a contradiction. Hence, $\frac{1}{2}(k-1) \leqq s_k$.

The second inequality is dual to the first. In the converse tournament T' with score sequence $t_1 \leqq t_2 \leqq \cdots \leqq t_p$ in which

$$t_i = (p-1) - s_{p-i+1},\ t_{p-k+1} \geqq \tfrac{1}{2}[(p-k+1)-1] = \tfrac{1}{2}(p-k),$$

by the first inequality of the theorem. But $s_k = (p-1) - t_{p-k+1}$, so

$$s_k \leqq (p-1) - \tfrac{1}{2}(p-k) = \tfrac{1}{2}(p+k-2),$$

proving the result.

The next theorems and their corollaries provide information regarding the scores of points and strong components of tournaments.

THEOREM 11. *Let v_i and v_j be points in different strong components of a tournament T. Then the line v_iv_j is in T if and only if $s_i > s_j$.*

Let v_i and v_j be points in different strong components of T. Let the line v_iv_j be in T. There is no cycle containing them. Every point adjacent from v_j is also adjacent from v_i, for otherwise we have a cycle containing v_i and v_j. In addition, v_j is adjacent from v_i. Therefore, $s_i > s_j$. Conversely, if $s_i > s_j$, the line v_jv_i cannot occur in T by this same argument. Hence, since T is complete, the line v_iv_j is in T.

COROLLARY 11a. *In a tournament, any two points with the same score are in the same strong component.*

The next theorem gives a sufficient condition for points with unequal scores to be in the same strong component. Although it is not a necessary condition, it may often be used in determining the strong components of a tournament.

THEOREM 12. *Let T be a tournament with score sequence $s_1 \leqq s_2 \leqq \cdots \leqq s_p$. If $0 \leqq s_n - s_m < \frac{1}{2}(n-m+1)$, then v_m and v_n are in the same strong component.*

Suppose that there are points v_m and v_n with $0 \leqq s_n - s_m < \frac{1}{2}(n-m+1)$, but they are not in the same strong component. Let j be the greatest integer less than n such that v_j is not in $S(v_n)$, the strong component containing v_n. By Theorem 10 we know that the score of v_n in the subtournament $S(v_n)$is not less than $\frac{1}{2}(n-j-1)$. Also by Theorem 11, there is a line from v_n to each of the j points v_i with $s_i \leqq s_j$. Therefore,

$$s_n \leqq j + \tfrac{1}{2}(n - j - 1) = \tfrac{1}{2}(n + j - 1).$$

Using Theorem 11 again, we know that every point adjacent from v_m is in the set $\{v_1, v_2, \cdots, v_j\}$. Applying Theorem 10 to the subtournament whose points are $v_1, v_2, \cdots, v_j$, we see that $s_m \leqq \frac{1}{2}(j+m-2)$. Combining these results with the assumption, we have $\frac{1}{2}(n-m+1) > s_n - s_m \geqq \frac{1}{2}(n+j-1) - \frac{1}{2}(j+m-2) = \frac{1}{2}(n-m+1)$, which is a contradiction.

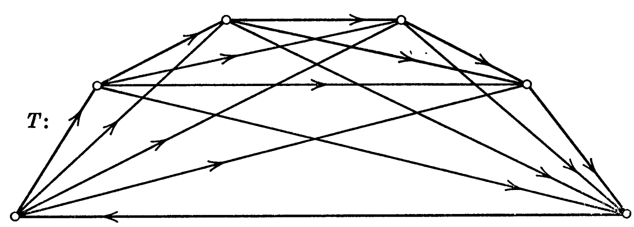


FIG. 7

We may illustrate Theorem 12 by referring again to Figure 7. From Corollary 11a we were able to conclude that v_1, v_2, v_3 all lie in the same strong component since the scores of these points are all 1. We see that $s_7 = 5$ and $s_4 = 4$. Substituting into the inequality of Theorem 12 we obtain $0 \leqq 5 - 4 = 1 < \frac{1}{2}(7-4+1) = 2$. Therefore, we conclude that v_4 and v_7 are in the strong component $S(v_4)$ and, *a fortiori*, so are v_5 and v_6.

The next two corollaries give sufficient conditions for a tournament to be strong.

COROLLARY 12a. *If the difference between every two scores in a tournament T is less than $\frac{1}{2}p$, then T is strong.*

Since the scores are $s_1 \leqq s_2 \leqq \cdots \leqq s_p$, the greatest difference between any two scores is $s_p - s_1$. By hypothesis, $s_p - s_1 < \frac{1}{2}p$ so that s_p and s_1 are in the same strong component by Theorem 12. Similarly for any other point s_i, it follows

from the theorem that s_i and s_1 are in the same strong component. Thus the entire tournament is strong.

COROLLARY 12b. *If both the outdegree and indegree of each point of a tournament T is at least $\frac{1}{4}(p-1)$, then T is strong.*

If id $v \geqq \frac{1}{4}(p-1)$, then

$$\text{od}\, v = p - 1 - \text{id}\, v \leqq p - 1 - \tfrac{1}{4}(p-1) = \tfrac{3}{4}(p-1).$$

Thus $s_1 \geqq \frac{1}{4}(p-1)$ and $s_p \leqq \frac{3}{4}(p-1)$. Hence

$$s_p - s_1 \leqq \tfrac{3}{4}(p-1) - \tfrac{1}{4}(p-1) = \tfrac{1}{4}(2p-2) < \tfrac{1}{2}p.$$

Therefore, by the preceding corollary, T is strong.

The strong tournament T displayed in Figure 7 shows that the conditions of these corollaries are not necessary for a tournament to be strong. In T, the greatest and least scores are 4 and 1, so that $s_p - s_1 = 3 = \frac{1}{2}p$. Also, the least score is less than $\frac{1}{4}(p-1)$. Nevertheless, the criteria are so simple that they sometimes are of considerable value in determining strong tournaments.

Let T_p be the number of distinct (non-isomorphic) tournaments with p points and let S_p be the number of strong tournaments among these. An explicit formula has been found for T_p by R. L. Davis [19]. The same formula is obtained in [20] from the number of oriented graphs by taking those oriented graphs with p points having $\frac{1}{2}p(p-1)$ lines. Thus the generating function

$$T(x) = \sum_{p=1}^{\infty} T_p\, x^p$$

may be regarded as known. J. W. Moon has pointed out that it is very easy to determine the corresponding generating function $S(x)$ in terms of $T(x)$ by using Corollary 5b. First write $T(x) = \sum_{n=1}^{\infty} T_n(x)$, where $T_n(x)$ counts those tournaments having exactly n strong components. Then substitute $T_n(x) = [S(x)]^n$ to obtain

$$S(x) = \frac{T(x)}{1 + T(x)}\,.$$

This article was written while the first named author was on sabbatical leave during 1962–63 at University College, London, with support from the National Science Foundation under grant NSF-GP 207. It is the basis for the chapter on tournaments in the book by Harary, Norman, and Cartwright [9]. We thank L. W. Beineke for his assistance in preparing this manuscript.

References

1. K. J. Arrow, Social Choice and Individual Values, New York, 1951.

2. L. W. Beineke and F. Harary, The maximum number of strong subtournaments, Canadian Math. Bull., 8(1965)491-498.

3. C. Berge, Théorie des graphes et ses applications, Paris, 1958. English translation, Wiley, New York, 1960.

4. D. Black, The Theory of Committees and Elections, Cambridge University Press, 1958.

5. P. Camion, Chemins et circuits hamiltoniens des graphes complets, Comptes Rendus. 249 (1959) 2151–2152.

6. L. R. Ford, Jr. and D. R. Fulkerson, Flows in Networks, Princeton University Press, Princeton, N. J., 1962.

7. J. D. Foulkes, Directed graphs and assembly schedules. Proc. Symp. Applied Math., Combinatorial Analysis, 10 (1960) 281–289.

8. D. R. Fulkerson, Upsets in round robin tournaments, RAND Memorandum RM-4112-PR, June 1963.

9. F. Harary, R. Norman, and D. Cartwright, Structural Models: an Introduction to the Theory of Directed Graphs, Wiley, New York, 1965.

10. J. G. Kemeny, J. L. Snell, and G. L. Thompson, Introduction to Finite Mathematics, Prentice-Hall, Englewood Cliffs, 1957, p. 310.

11. M. G. Kendall and B. Babington Smith, On the method of paired comparisons, Biometrika, 31 (1940) 324–325.

12. D. König, Theorie der endlichen und unendlichen Graphen, Leipzig, 1936. Chelsea, New York, 1950.

13. H. G. Landau, On dominance relations and the structure of animal societies: III. The condition for a score structure, Bull. Math. Biophysics, 15 (1953) 143–148.

14. R. D. Luce and H. Raiffa, Games and Decisions, Wiley, New York, 1957.

15. J. W. Moon, An extension of Landau's theorem on tournaments, Pacific J. Math., 13 (1963) 1343–1347.

16. O. Ore, Graphs and their Uses, New York, 1963.

17. L. Rédei, Ein kombinatorischer Satz, Acta Litterarum ac Scientiarum (Sectio Scientiarum Mathematicarum) Szeged, 1 (1934) 39–43.

18. D. Silverman, Problem 463, Math. Mag., 35 (1962) 189.

19. R. L. Davis, Structures of dominance relations, Bull. Math. Biophysics, 16 (1954) 131–140.

20. F. Harary, The number of oriented graphs, Mich. Math. J., 4 (1957) 221–224.

TRANSLATIONS ON AN ABSTRACT SPACE

HYMAN GABAI, University of Illinois and UICSM

Introduction. Let $\mathcal{E}$ be a nonempty set, and let $\mathfrak{I}$ be such a partition of $\mathcal{E}\times\mathcal{E}$ that each member of $\mathfrak{I}$ is a function of $\mathcal{E}$ into $\mathcal{E}$. For any $(A, B)\in\mathcal{E}\times\mathcal{E}$, (A, B) is contained in only one member of the partition of $\mathcal{E}\times\mathcal{E}$, and so each member of $\mathfrak{I}$ is determined by a point of $\mathcal{E}$ and its image. The function in $\mathfrak{I}$ which contains (A, B) will be denoted by $B-A$. The image of $X\in\mathcal{E}$ under the mapping $B-A$ will be denoted by $(B-A)+X$. We shall also use '+' for function composition, and we define:

$$(A - B) + (C - D) = (A - B) \circ (C - D).$$

We also assume that the members of $\mathfrak{I}$ satisfy:

$$A - A = B - C \Leftrightarrow B = C, \tag{1}$$

$$(A - B) + (B - C) = A - C. \tag{2}$$

We first notice that $(B-A)+A=B$ and $[(B-A)+C]-C=B-A$.

In this paper we shall investigate some of the properties of $\mathfrak{I}$ and some of the relationships between the set $\mathcal{E}$ and the set $\mathfrak{I}$.

Example 1. If $\mathcal{E}=\{a\}$, then $\mathfrak{I}$ is just the identity mapping on $\mathcal{E}$.

Example 2. If $\mathcal{E}=R^n$, let $\mathfrak{I}=\{f_a: a\in R^n\}$, where $f_a(x)=x+a$.

Example 3. If $(\mathcal{E}, +, 0, -)$ is a group, let $\mathfrak{I}=\{f_A: A\in\mathcal{E}\}$, where $f_A(X)=X+A$. The set $\mathfrak{I}$ forms a group isomorphic with $\mathcal{E}$. [$\mathfrak{I}$ is a subgroup of the group of all one-to-one transformations of $\mathcal{E}$ onto itself.]

Example 4. If $\mathcal{E}=R$, let $\mathfrak{I}=\{g_a: a\in R\}$, where we define:

$$[\text{if } a \text{ is rational}] \quad g_a(x) = \begin{cases} x - a & [x \text{ is rational}] \\ x + a & [x \text{ is irrational}] \end{cases}$$

$$[\text{if } a \text{ is irrational}] \quad g_a(x) = x + a.$$

Example 5. Let $\mathfrak{I}$ be a partition of $\mathcal{E}\times\mathcal{E}$ as described in the introduction and let h be any one-to-one transformation of $\mathcal{E}$ onto itself. Then the set $\mathfrak{I}'=\{h^{-1}\circ f\circ h: f\in\mathfrak{I}\}$ also satisfies the conditions set forth in the introduction.

THEOREM 1. *The quadruple* $(\mathfrak{I}, +, A-A, -)$ *is a group of one-to-one transformations of* $\mathcal{E}$ *onto itself, where we define* $-(B-A)=A-B$.

Proof. $\mathfrak{I}$ is closed under addition, because for any transformations $(A-B)$ and $(C-D)$ we may write: $A-B=[(A-B)+C]-C$ and we know that

$$\{[(A-B)+C]-C\}+(C-D)\in\mathfrak{I}.$$

The identity element in $\mathfrak{I}$ is $A-A$, because for any transformation $B-C$:

$$\begin{aligned}(B-C)+(A-A) &= \{[(B-C)+A]-A\}+(A-A) \\ &= [(B-C)+A]-A = B-C.\end{aligned}$$

The inverse of $B-A$ is $A-B$ since $(B-A)+(A-B)=B-B$. The addition operation in $\mathfrak{I}$ is associative since it is function composition. Therefore $\mathfrak{I}$ is a group.

For future use we also note that $(A-A)$ is the identity mapping on $\mathcal{E}$. Because for any $X\in\mathcal{E}$:

$$[(A-A)+X]-X = A-A$$

and this implies that $(A-A)+X=X$.

Each member of $\mathfrak{I}$ is a one-to-one mapping since if $(B-A)+X=(B-A)+Y$ then:

$$\begin{aligned}X &= (A-A)+X = (A-B)+[(B-A)+X] \\ &= (A-B)+[(B-A)+Y] = Y.\end{aligned}$$

Each member of $\mathfrak{I}$ is an onto mapping since for any $(B-A)\in\mathfrak{I}$ and any $X\in\mathcal{E}$ we have:

$$X = (B-B)+X = (B-A)+[(A-B)+X].$$

THEOREM 2. *$\mathfrak{I}$ is a commutative group if and only if for any $(B-A)\in\mathfrak{I}$ and any points X, Y of $\mathcal{E}$: $[(B-A)+X]-[(B-A)+Y]=X-Y$.*

Proof. If $\mathfrak{I}$ is commutative then

$$\begin{aligned}&[(B - A) + X] - [(B - A) + Y]\\ &\quad = [(B - A) + \{(X - Y) + Y\}] - [(B - A) + Y]\\ &\quad = [(X - Y) + \{(B - A) + Y\}] - [(B - A) + Y] = X - Y.\end{aligned}$$

To prove the converse, let $B-A$ and $C-D$ be members of $\mathfrak{I}$. Then:

$$\begin{aligned}&(B - A) + (C - D)\\ &\quad = \{[(C - D) + B] - [(C - D) + A]\} + \{[(C - D) + A] - A\}\\ &\quad = [(C - D) + B] - A = \{[(C - D) + B] - B\} + (B - A)\\ &\quad = (C - D) + (B - A).\end{aligned}$$

REMARK 1. If $\mathfrak{I}$ is postulated to be commutative then we may omit the condition (1) given in the introduction:

$$A - A = B - C \Leftrightarrow B = C.$$

For if $B=C$ and $\mathfrak{I}$ is commutative then:

$$B - C = B - B = (B - A) + (A - B) = (A - B) + (B - A) = A - A.$$

Conversely, if $A-A=B-C$ and $\mathfrak{I}$ is commutative then:

$$\begin{aligned}B &= (B - C) + C = (A - A) + C = [(A - C) + (C - A)] + C\\ &= [(C - A) + (A - C)] + C = (C - A) + A = C.\end{aligned}$$

DEFINITION 1. *If $\mathfrak{I}$ is commutative then $\mathfrak{I}$ will be called a set of translations on $\mathcal{E}$.*

REMARK 2. The existence of a set of translations on the set of "points of space" is essentially equivalent to the first few postulates of a course in geometry currently being developed by UICSM. [See Remark 5 at the end of this paper.]

REMARK 3. All of the examples except Example 3 are examples of sets of translations. In Example 3, if $\mathcal{E}$ is a commutative group then $\mathfrak{I}$ is a set of translations on $\mathcal{E}$.

THEOREM 3. *Let $\mathfrak{I}$ and $\mathfrak{I}'$ be sets of translations on $\mathcal{E}$. For each $X\in\mathcal{E}$, let $(X-A)'\circ(A-B)=(A-B)\circ(X-A)'$, where $(A-B)$ and $(A-B)'$ are translations in $\mathfrak{I}$ and $\mathfrak{I}'$ respectively. Then $B-A=(B-A)'$.*

Proof. We denote the addition operation in $\mathfrak{I}$ and $\mathfrak{I}'$ by '$+$' and '$\oplus$' respectively. Also, $f * B$ will be the image of B under the function f. For any $X\in\mathcal{E}$:

$$\begin{aligned}(A - B) + [(B - A)' \oplus X] &= (A - B) + \{(B - A)' \oplus [(X - B)' \oplus B]\}\\ &= (A - B) + \{(X - A)' \oplus B\} = [(A - B) \circ (X - A)'] * B\end{aligned}$$

$$= [(X - A)' \circ (A - B)] * B = (X - A)' \oplus [(A - B) + B]$$
$$= (X - A)' \oplus A = X.$$

Therefore,

$$[(B - A) + (A - B)] + [(B - A)' \oplus X] = (B - A) + X,$$

and we conclude that $(B-A)' = (B-A)$.

THEOREM 4. *Let $\mathcal{E}$ be a metric space and $\mathfrak{T}$ a set of translations on $\mathcal{E}$, satisfying:*

$$(A, B) \in C - D \Rightarrow d(A, B) = d(C, D).$$

Then each translation in $\mathfrak{T}$ is a distance preserving homeomorphism of $\mathcal{E}$ onto $\mathcal{E}$.

Proof. We first notice that by our hypothesis we may define:

$$\|A - B\| = d(A, B).$$

Since $\mathfrak{T}$ is commutative we have for any $B-A \in \mathfrak{T}$ and any points X, Y of $\mathcal{E}$:

$$d(X, Y) = \|X - Y\| = \|[(B - A) + X] - [(B - A) + Y]\|$$
$$= d([(B - A) + X], [(B - A) + Y]).$$

Therefore $(B-A)$ is distance preserving and continuous. Since $-(B-A) = (A-B)$, the inverse is also continuous. We have already seen that $B-A$ is one-to-one and onto, so $(B-A)$ is a homeomorphism.

THEOREM 5. *Let $\mathcal{E}$ be a topological space and $\mathfrak{T}$ a set of translations on $\mathcal{E}$, satisfying the condition that if N_A is a neighborhood of A then $(B-A)+N_A$ is a neighborhood of B. Then each translation in $\mathfrak{T}$ is a homeomorphism of $\mathcal{E}$ onto $\mathcal{E}$.*

Proof. We have already seen that each translation is one-to-one and onto, so we need only show that it is continuous. Let $X \in \mathcal{E}$ and $N_{(A-B)+X}$ a neighborhood of $(A-B)+X$. Now the inverse of $(A-B)$ is $(B-A)$ and:

$$B - A = \{(B - A) + [(A - B) + X]\} - [(A - B) + X]$$
$$= X - [(A - B) + X].$$

Therefore the inverse image of $N_{(A-B)+X}$ is

$$(B - A) + N_{(A-B)+X} = \{X - [(A - B) + X]\} + N_{(A-B)+X}$$

which is a neighborhood of X. Therefore $(A-B)$ is continuous. Since $(B-A)$ is also continuous, $(A-B)$ is a homeomorphism.

In the next theorem we investigate the natural topology induced on $\mathfrak{T}$ by the topology on $\mathcal{E}$. We recall that $\mathfrak{T}$ is the set of equivalence classes of a partition on $\mathcal{E} \times \mathcal{E}$.

THEOREM 6. *Let $\mathcal{E}$ be a topological space, and $\mathfrak{T}$ be endowed with the quotient topology with respect to the product topology on $\mathcal{E} \times \mathcal{E}$. Then $\mathfrak{T}$ is homeomorphic to $\mathcal{E}$, and for each $O \in \mathcal{E}$ a homeomorphism h_o can be defined.*

Proof. Let $O\in\mathcal{E}$. Then $\mathfrak{T}=\{(A-O): A\in\mathcal{E}\}$. Now $\{O\}\times\mathcal{E}\subseteq\mathcal{E}\times\mathcal{E}$, and the mappings

$$\begin{gathered}\mathcal{E}\to\{O\}\times\mathcal{E}\to\mathfrak{T}\\ X\to(O, X)\to(X-O)\end{gathered}$$

are both homeomorphisms [by the definition of the topologies on $\mathcal{E}\times\mathcal{E}$, $\{O\}\times\mathcal{E}$, and $\mathfrak{T}$]. These mappings define the homeomorphism h_o.

DEFINITION 2. *If $\mathcal{E}$ is a topological space, the topology on $\mathfrak{T}$ described in Theorem 6 will be called the natural topology on $\mathfrak{T}$. If, further, the condition of Theorem 5 is satisfied so that each translation in $\mathfrak{T}$ is a homeomorphism of $\mathcal{E}$ onto $\mathcal{E}$, then the pair $(\mathcal{E}, \mathfrak{T})$ will be called a translation space.*

THEOREM 7. *Let $(\mathcal{E}, \mathfrak{T})$ be a translation space such that for each $A\in\mathcal{E}$ and each neighborhood U_A of A, there exists a neighborhood V_A of A such that:*

$$P\in V_A \text{ and } Q\in V_A\Rightarrow(Q-A)+P\in U_A.$$

Then $\mathfrak{T}$ is a topological group.

Proof. $\mathfrak{T}$ is a topological space and a commutative group. We first show that $+$ is continuous. If $O\in\mathcal{E}$ then $\mathfrak{T}=\{(A-O): A\in\mathcal{E}\}$, and the mapping

$$\begin{gathered}\mathcal{E}\to\mathfrak{T}\\ A\to(A-O)\end{gathered}$$

is a homeomorphism. So instead of considering the mapping:

$$\begin{gathered}\mathfrak{T}\times\mathfrak{T}\to\mathfrak{T}\\ ((A-O), (B-O))\to(A-O)+(B-O)\end{gathered}$$

we notice that:

$$(A-O)+(B-O)=\{[(A-O)+(B-O)]+O\}-O=\{(B-O)+A\}-O$$

and consider the mapping:

$$\begin{gathered}\mathcal{E}\times\mathcal{E}\to\mathcal{E}\\ (A, B)\to(B-O)+A.\end{gathered}$$

Now let U be any neighborhood of $(B-O)+A$. Then by the hypothesis, there exists a neighborhood V of $(B-O)+A$ such that

$$P\in V \text{ and } Q\in V\Rightarrow\{Q-[(B-O)+A]\}+P\in U.$$

If W_A is any neighborhood of A then

$$(B-O)+A\in(B-O)+W_A$$

and since $(B-O)$ is a homeomorphism, we may choose W_A so that $(B-O)+W_A\subseteq V$. In the same way we may choose W_B a neighborhood of B so that

$$(B - O) + A = (A - O) + B \in (A - O) + W_B \subseteq V.$$

Then for any $(P, Q) \in W_A \times W_B$, $(B-O)+P \in V$ and $(A-O)+Q \in V$, and therefore:

$$(Q - O) + P = \{[(A - O) + Q] - [(B - O) + A]\} + [(B - O) + P] \in U.$$

So, $+$ is continuous.

To see that the function

$$\mathfrak{T} \to \mathfrak{T}$$
$$(A - O) \to -(A - O)$$

is continuous, we notice that

$$-(A - O) = \{[(O - A) + (O - A)] + A\} - O$$

and that the function

$$\mathcal{E} \to \mathcal{E}$$
$$A \to [(O - A) + (O - A)] + A$$

is continuous.

THEOREM 8. *Let the hypothesis of Theorem 4 be satisfied, and let $\mathfrak{T}$ be endowed with its natural topology. [Therefore $(\mathcal{E}, \mathfrak{T})$ is a translation space.] Let multiplication of translations by scalars be defined so that $\mathfrak{T}$ is a vector space. If $O \in \mathcal{E}$, then $\mathfrak{T} = \{(A-O): A \in \mathcal{E}\}$ and we use the notation:*

$$(A - O)a = (A_a - O).$$

We further assume that the multiplication satisfies: $d(A_a, B_a) = |a| d(A, B)$. Then $\mathfrak{T}$ is a normed topological vector space.

Proof. As in Theorem 4, define $d(A, B) = ||A - B||$. Then the norm of $(A-B)$ in $||A-B||$ since:

$$\begin{aligned} &||(A - B) + (C - D)|| \\ &\quad = ||A - [(D - C) + B]|| \\ &\quad = d(A, (D - C) + B) \leqq d(A, B) + d(B, (D - C) + B) \\ &\quad = ||A - B|| + ||B - [(D - C) + B]|| = ||A - B|| + ||C - D|| \end{aligned}$$

and:

$$\begin{aligned} ||(A - B)a|| &= ||(A - O)a - (B - O)a|| = ||A_a - B_a|| \\ &= d(A_a, B_a) = |a| d(A, B) = |a| ||A - B||. \end{aligned}$$

Since the hypothesis of this theorem satisfies Theorem 7, in order to show that $\mathfrak{T}$ is a topological vector space we need only prove that the multiplication of translations by scalars is continuous. Let $O \in \mathcal{E}$, and $\mathfrak{T} = \{(A-O): A \in \mathcal{E}\}$.

Consider the mappings:

$$\mathcal{E} \times R \to \mathfrak{I} \times R \to \mathfrak{I} \to \mathcal{E}$$
$$(A, a) \to (A - O, a) \to (A - O)a \to A_a.$$

The first and last mappings are homeomorphisms, so to show that the second mapping is continuous, we may consider the mapping:

$$\mathcal{E} \times R \to \mathcal{E}$$
$$(A, a) \to A_a.$$

Let $A_a \in \mathcal{E}$ and $\epsilon > 0$ be given. Let $B \in \mathcal{E}$ such that $d(A, B) < \delta$ and let $b \in R$ such that $|a - b| < \eta$, for some $\delta > 0$ and $\eta > 0$. Let $b = a + \eta'$, where $|\eta'| < \eta$. Then, using the vector space properties of $\mathfrak{I}$, we may write:

$$\begin{aligned} d(A_a, B_b) &= \|A_a - B_b\| = \|(A - O)a - (B - O)b\| \\ &= \|(A - O)a - [(B - O)a + (B - O)\eta']\| \\ &= \|[(A - O) - (B - O)]a - (B - O)\eta'\| \\ &\leqq \|A - B\| \, |a| + \|B - O\| \, |\eta'| \\ &< \delta |a| + \{\|B - A\| + \|A - O\|\} |\eta'| \\ &< \delta |a| + \delta |\eta'| + \|A - O\| \, |\eta'|. \end{aligned}$$

For suitable choice of δ and η, this last sum can be made less than ϵ. Therefore the mapping is continuous, and $\mathfrak{I}$ is a topological vector space.

REMARK 4. If the hypothesis of Theorem 8 is satisfied, we know that $\mathfrak{I}$ is a normed topological vector space. In a natural way we may define addition of elements of $\mathcal{E}$, and multiplication of elements of $\mathcal{E}$ by scalars:

$$A + B = C \Leftrightarrow (A - O) + (B - O) = C - O$$
$$Aa = A_a \Leftrightarrow (A - O)a = (A_a - O).$$

Then $\mathcal{E}$ is also a normed topological vector space with $\|A\| = \|A - O\|$. The homeomorphism h_o:

$$\mathcal{E} \to \mathfrak{I}$$
$$A \to A - O$$

satisfies:

$$h_o(A + B) = h_o((A - O) + B) = (A - O) + (B - O) = h_o(A) + h_o(B)$$
$$h_o(Aa) = h_o(A_a) = (A - O)a = [h_o(A)]a.$$

Therefore h_o is a linear operator of $\mathcal{E}$ onto $\mathfrak{I}$, and so $\mathcal{E}$ and $\mathfrak{I}$ are topologically isomorphic.

REMARK 5. The University of Illinois Committee on School Mathematics [under a grant from the National Science Foundation] has prepared a course in

3-dimensional geometry using a vectorial approach. This course was written by Professor Herbert E. Vaughan, and it is currently being taught to tenth grade students at the University High School. In this vector development, we deal with three sets:

the set $\mathcal{E}$ of points of 3-dimensional euclidean space
the set $\mathfrak{T}$ of translations on $\mathcal{E}$
the set R of real numbers.

The postulates are introduced gradually, only after the students have been properly motivated by exercises and discussions. The first few postulates which are introduced are essentially equivalent to the statement that there exists a set of translations on the set of "points of space." Eventually $\mathfrak{T}$ is postulated to be a vector space, a three-dimensional vector space, and, finally, a three-dimensional inner product space.

In this paper we used Vaughan's notation to deal with the set $\mathfrak{T}$ since it permitted a convenient way to operate with these functions algebraically in a natural manner.

The author wishes to take this opportunity to thank Professors Vaughan and Beberman for their many informative discussions, both mathematical and pedagogical.

This work was supported by National Science Foundation Grant G23554.

References

1. D. Bushaw, Elements of General Topology, Wiley, New York, 1963.
2. Angus E. Taylor, Introduction to Functional Analysis, Wiley, New York, 1958.
3. University of Illinois Committee on School Mathematics (UICSM), A Vector Approach to Euclidean Geometry (in preparation).

UNIQUE MINIMAL REPRESENTATIONS WITH POSITIVE BASES

J. R. REAY, Western Washington State College

The set B positively spans the linear space L if each point of L may be represented as a linear combination of the points of B using only positive coefficients. If no proper subset of B has this property, then we say B is a *positive basis* for L. The theory of positive bases has been motivated by certain developments in linear programming and game theory, and by a rather natural usefulness when considering Helley-type convexity theorems in finite dimensional linear spaces, (see [4]). The theory is somewhat parallel to the ideas concerning linear bases for linear spaces, the notable exception being that various equivalent notions of linear independence have nonequivalent analogs when considering positive bases. In particular, it is not even true that a minimal representation of a point of L (i.e. linear representations with positive coefficients involving a minimal number of elements of B) in terms of B is unique.

The purpose of this note is to establish conditions equivalent to the uniqueness of a minimal representation (sec. 2), and point out (sec. 3) how certain machinery from the theory gives well-defined continuous representations in terms of an arbitrary positive basis B.

1. Introduction. We first establish necessary notation and give the standard definitions in a way that shows the similarity between the linear and positive structures.

For any subset B of the real linear space L, let $\mathcal{L}B$ denote the set of all real valued functions on B which have value zero at all but a finite number of points of B. Let $\mathcal{P}B$ be the set of all functions in $\mathcal{L}B$ which are everywhere nonnegative. A point $\sum_{b\in B}\alpha_b b\in L$ is called a *linear* [respectively, *positive*] *combination from* B if $\alpha\in\mathcal{L}B$ [respectively, $\alpha\in\mathcal{P}B$], and the collection of all such points is called the *linear* [*positive*] *hull* of B which we denote by lin B [pos B]. A set B *linearly* [*positively*] *spans* L if $L=\text{lin } B$ [$L=\text{pos } B$]. The set B is *linearly* [*positively*] *independent* if for each $b_0\in B$ it is true that, respectively,

$$b_0 \notin \text{lin}\,(B \sim \{b_0\}),\ [b_0 \notin \text{pos}\,(B \sim \{b_0\})].$$

The set B is a *linear* [*positive*] *basis* for L if it linearly [positively] spans L and is linearly [positively] independent.

An element α of $\mathcal{L}B[\mathcal{P}B]$, where B is any subset of L, is a *linear* [*positive*] *representation* of $x\in L$ in terms of B if $x=\sum_{b\in B}\alpha_b b$, and any representation α of 0 is called a *relation*. While the linear independence of the set B is equivalent to the property that there are no nontrivial linear relations, it is obvious that every positive basis has nontrivial positive relations. (Consider the sum of any positive representations of $x\neq 0$ and $-x$.) The *length* of a positive representation α of the point $x\in L$ in terms of B is the cardinality of the set $\{b\in B \mid \alpha_b\neq 0\}$, and α is called a *minimal representation* of x if no other representation of x (in terms of the set B) has shorter length. There are many positive bases for which a minimal representation of a point in L is unique. For example, if B is a linear basis for L, then $B\cup(-B)$ is clearly such a positive basis, and if L is also finite dimensional, then $B\cup\{\sum_{b\in B}-b\}$ is such a positive basis. ([**3**], Theorem 4.4.) The set $B=\{(1, 0, 0), (-1, 1, 0), (-1, -1, 0), (-1, 0, 1), (-1, 0, -1)\}$ is a positive basis for E^3, but $(-1, 0, 0)$ does not have a unique minimal representation.

Let B be a positive basis for L. It is clear that not all subsets of B form positive bases for linear subspaces of L. We say that a linear subspace L' of L is a *spanned subspace with respect to* B if $(L'\cap B)$ is a positive basis for L'. If, moreover, pos $(L'\cap B)$ has finite dimension n, and card $(L'\cap B)=n+1$, then we say L' is a *minimal subspace* (with respect to B) and $(L'\cap B)$ is a *minimal positive basis* for L'. As the above example shows, not all spanned subspaces are minimal subspaces (E^3 is spanned, but not minimal). But, if $b_0\in B$ and B' is a subset of B which appears in a minimal representation α of $-b_0$ (i.e., $\alpha_b>0$ iff $b\in B'$),

then it follows that $\{b_0\} \cup B'$ is a minimal positive basis for the linear space that it spans. (See Theorem 3.2 of [1].) Thus each $b \in B$ lies in some minimal subspace, and L is the linear sum of its minimal subspaces. Note that B is a minimal positive basis for L if and only if for each $b \in B$, $B \sim \{b\}$ is a linear basis for L.

McKinney [3] calls the positive basis B a *strong positive basis* if $\text{pos } B_1 \cap \text{pos } B_2 \subseteq \{0\}$ whenever B_1, B_2 are disjoint subsets of B. In view of the fact that minimal representations in terms of a minimal positive bases are unique ([3], Theorem 4.4), it follows that if L is the direct linear sum of its minimal subspaces with respect to B then B is a strong positive basis.

2. Equivalences of unique minimal representations. Let B be a positive basis for the linear space L. We will consider the following conditions on B.

(a) The minimal representation of each $x \in L$ (in terms of B) is unique.
(b) B is a strong positive basis.
(c) Distinct minimal subspaces have only $\{0\}$ in common.
(d) Distinct minimal positive bases for minimal subspaces are pair-wise disjoint.
(e) L is the direct linear sum of its minimal subspaces.
(f) B admits a unique partitioning into pair-wise disjoint subsets, each of which is a minimal positive basis for a minimal subspace of L.
(g) B admits a partitioning into pair-wise disjoint subsets, each of which is a minimal positive basis for a minimal subspace of L.

THEOREM 1. *Let n denote the dimension of L. The following implications hold between the above conditions on B:*

$$\text{a} \leftarrow \text{b} \leftrightarrow \text{c} \leftrightarrow \text{d} \leftrightarrow \text{e} \rightarrow \text{f} \rightarrow \text{g} \quad \textit{for all } n$$

$$\text{a} \leftrightarrow \text{b} \quad \textit{if } n \textit{ is finite}$$

$$\text{e} \leftrightarrow \text{f} \quad \textit{if } n \leqq 5$$

$$\text{f} \leftrightarrow \text{g} \quad \textit{if } n \leqq 4.$$

Proof. Essentially, only the equivalences involving conditions d and f are new. The implications b←e→c→d→f→g in spaces of arbitrary dimension are either obvious or straight forward. McKinney [3] establishes the implications b→c→(a and e) for spaces of arbitrary dimension, (pages 144–5), and Bonnice-Klee [1] establish the implication a→c for the case where B is finite. In view of the fact that $n+1 \leqq \text{card } B \leqq 2n$ in an n-dimensional space L for all positive bases B (Theorems 3.8 and 6.7 of Davis [2]), the implication a→c holds if L is finite dimensional. (Bonnice-Klee also give an example to show a→c fails if L is $\aleph_1$-dimensional, but the problem in an $\aleph_0$-dimensional space is unfortunately still open.)

We will complete the proof by showing that g→d if $n \leqq 4$, f→d if $n \leqq 5$ and finally d→c if dim $L = n$ is arbitrary. This is done in the following lemmas.

LEMMA 2. *Suppose the following conditions hold in a linear space L of arbitrary dimension.*

(1) *B is a positive basis for L.*
(2) *$B_1 \subset B$ is a minimal positive basis for a minimal subspace of dimension d_1.*
(3) *$B_2 \subset B$, $B_1 \cap B_2 = \phi$, and* pos B_2 *is a linear subspace of L.*

Then the dimension d of the linear subspace $M = (\text{pos } B_1 \cap \text{pos } B_2)$ is at most $\max(0, d_1 - 2)$.

Proof. Note that $M \neq \text{pos } B_1$ or else B is not positively independent. Thus $d < d_1$ and the denial of the lemma asserts that $1 \leqq d = d_1 - 1$. This asserts that M is a hyperplane through the origin in pos B_1. Now some element x of M must have length d_1 (in terms of B_1), for otherwise the points of B_1 would form the vertices of a simplex and M would be a hyperplane which met the interior of the simplex but failed to meet the relative interior of any $(d_1 - 1)$-dimensional face of the simplex. Thus suppose x is a point of M whose minimal representation

$$x = \sum_{b \in (B_1 \sim \{b_0\})} \alpha_b b$$

has length d_1, i.e., each α_b is strictly positive. Then $\{-x\} \cup (B_1 \sim \{b_0\})$ forms a positive basis for pos B_1. Thus

$$b_0 \in \text{pos } (\{-x\} \cup (B_1 \sim \{b_0\})) \subset \text{pos } (B_2 \cup (B_1 \sim \{b_0\})) = \text{pos } (B \sim \{b_0\}).$$

Hence B is not positively independent, a contradiction.

LEMMA 3. *If the dimension n of L is at most* 4, *then condition* g *is equivalent to condition* d.

Proof. The implication g→d when $n \leqq 4$ follows easily from several results of Davis [2]. Suppose $\dim L = n \leqq 4$ and condition g holds, i.e., $B = B_1 \cup \cdots \cup B_k$ is a partitioning of B such that each B_i is a minimal positive basis for the linear space pos B_i. We may assume $n = 4$, whence $5 \leqq \text{card } B \leqq 8$. The cases card $B = 5$ or 8 are easily handled ([4], Theorem 1.3). Letting m denote the number of distinct minimal positive bases contained in B, it may be shown that card $B = n + m$ whenever $n \leqq 4$ ([2], Theorems 6.6 and 6.8). If card $B = 6$ then $m = 2$. Since the partition must involve at least 2 minimal positive bases, it follows that $k = 2$, i.e., $B = B_1 \cup B_2$ and distinct minimal bases must be pair-wise disjoint. The only remaining case is when card $B = 7$, $n = 4$, $m = 3$, and $2 \leqq k \leqq 3$. If $k = 3$, i.e., $B = B_1 \cup B_2 \cup B_3$, then all minimal positive bases occur in the partition since $m = 3$, and condition d follows. Thus assume $k = 2$, i.e., $B = B_1 \cup B_2$. The same result of Davis shows that $(\text{pos } B_1) \cap (\text{pos } B_2) = \{0\}$. Thus the dimensions of these linear subspaces must be 3 and 1 or else 2 and 2, and in either case card $(B_1 \cup B_2) = 6$. But $\text{card}(B_1 \cup B_2) = \text{card } B = 7$. Thus this case cannot occur. This shows g→d when $n \leqq 4$.

LEMMA 4. *If the dimension n of L is at most* 5, *then condition* f *is equivalent to condition* d.

Proof. To show f$\rightarrow$d if $n \leqq 5$, it suffices to consider only the case where $n=5$. Let B be a positive basis for the 5-dimensional linear space L, and let $B = B_1 \cup \cdots \cup B_k$ be a unique partition of B into minimal positive bases for minimal subspaces. Then $6 \leqq \text{card } B \leqq 10$ and therefore $1 \leqq k \leqq 5$. If $k=1$ or $k=5$, then B is respectively a minimal or a maximal basis for L and each minimal positive basis appears in the partition of B. Thus, distinct minimal positive bases are pair-wise disjoint and condition d holds. Hence we consider only the cases where $2 \leqq k \leqq 4$. This implies $7 \leqq \text{card } B \leqq 9$ ([4], Theorem 1.3).

Case 1. $k=2$, $B = B_1 \cup B_2$. Let d_1, d_2 and d denote respectively the dimensions of the linear spaces pos B_1, pos B_2, and (pos B_1)$\cap$(pos B_2). We may assume that $0 \leqq d < d_1 \leqq d_2$. Suppose $B_3 \subset B$ is any minimal positive basis distinct from B_1 and B_2. The set B_3 must have elements in common with both B_1 and B_2, but B_3 does not contain B_1 or B_2.

Case 1.1. Assume that $d=0$. Then $L = (\text{pos } B_1) \oplus (\text{pos } B_2)$ and $d_1 + d_2 = 5$. If a distinct minimal basis B_3 exists, then (pos B_1)$\cap$(pos B_3) is of dimension at least one and therefore $d_1 \geqq 2$. Hence we may assume $d_1 = 2$, $d_2 = 3$, and the linear space (pos B_1)$\cap$(pos B_3) must be one-dimensional. Let b_0 be the element of $B_1 \cap B_3$. Note that $-b_0 \in \text{pos } B_3$ but $-b_0 \notin \text{pos } (B_1 \cap B_3) = \text{pos } \{b_0\}$. The element $-b_0$ has a minimal positive representation

$$-b_0 = \sum_{b \in B_3} \alpha_b b = \sum_{b \in (B_3 \cap B_1)} \alpha_b b + \sum_{b \in (B_3 \cap B_2)} \alpha_b b.$$

But since $-b_0 \in \text{pos } B_1$ and L is the direct linear sum of pos B_1 and pos B_2, it follows that $\alpha_b = 0$ for all $b \in (B_3 \cap B_2)$. Thus

$$-b_0 = \sum_{b \in (B_3 \cap B_1)} \alpha_b b$$

and $-b_0 \in \text{pos } (B_1 \cap B_3)$, a contradiction. Thus no such set B_3 exists and again condition d of the theorem is valid.

Case 1.2. Assume $d=1$. Then Lemma 2 asserts that $d_1 \geqq 3$. Since $n=5$, it follows that $d_1 = d_2 = 3$. That is, L is the sum of the two 3-dimensional spaces pos B_1 and pos B_2 whose intersection $M = \text{pos } B_1 \cap \text{pos } B_2$ is 1-dimensional. If $x \in M$ and $x \neq 0$, it is clear that the length of a minimal representation of x (in terms of either B_1 or B_2) is at least 2. Suppose

$$x = \sum_{b \in (B_1 \sim \{b_0\})} \alpha_b b$$

is a minimal representation of x of length 3. Then the same argument used in Lemma 2 shows that $b_0 \in \text{pos } (B \sim \{b_0\})$ and B is not positively independent. We may thus assume that the length of α is exactly 2. Now let α_i be a minimal

representation of x in terms of B_i, and let β_i be a minimal representation of $-x$ in terms of B_i, $i=1, 2$. Let $C_i=\{b\in B_i \mid b$ appears in $\alpha_i\}$, and let $D_i=\{b\in B_i \mid b$ appears in $\beta_i\}$, $i=1, 2$. (These sets are uniquely determined because B_i is a minimal positive basis.) Then $B_i=C_i\cup D_i$ must be a partitioning of B_i into two disjoint subsets, each of cardinality 2. It is also clear that $B_3=C_1\cup D_2$ and $B_4=C_2\cup D_1$ are minimal positive bases for 3-dimensional minimal subspaces of L. Further $B=B_3\cup B_4$ is a partitioning of B. Thus condition f fails because the partitioning is not unique.

Case 1.3. Assume $d\geqq 2$. Lemma 2 asserts that $d_1\geqq 4$. Thus card $B_2\geqq$card B_1 $\geqq 5$ and card $B\geqq 10$, a contradiction.

Case 2. $k=3$. $B=B_1\cup B_2\cup B_3$ is a unique partitioning of B into minimal positive bases for minimal subspaces. Assume without loss of generality that card B_1 $\leqq$card $B_2\leqq$card B_3. Since card $B\leqq 9$, it follows that card $B_1\leqq 3$ and pos B_1 is at most 2-dimensional. Thus the dimension d of the linear space $M=$pos B_1 $\cap$pos $(B_2\cup B_3)$ is zero by Lemma 2. This implies that L is the direct sum of pos B_1 and pos $(B_2\cup B_3)$. Now assume $B_4\subset B$ is any minimal positive basis distinct from B_i, $i=1, 2, 3$. Then $B_1\cap B_4$ has at most one element, say b_0. As before, it follows that $-b_0\in$pos $(B_4\sim\{b_0\})\subset$pos $(B_2\cup B_3)$ denying $L=$pos $B_1\oplus$pos $(B_2\cup B_3)$. Thus $B_4\subset B_2\cup B_3$. But dim $(B_2\cup B_3)\leqq 4$ and the existence of B_4 denies Lemma 3. Thus no such set B_4 can exist and the distinct minimal positive bases of B are pair-wise disjoint.

Case 3. $k=4$, $B=B_1\cup B_2\cup B_3\cup B_4$. As in the last case, we may assume card $B_1\leqq 3$, and the proof proceeds as above. This completes the proof of the lemma.

LEMMA 5. *If the dimension n of L is arbitrary, then condition* d *is equivalent to condition* c.

Proof. Let C and D be distinct subsets of B which are minimal positive bases for the minimal subspaces pos C and pos D, and suppose that there exists a nonzero point $x\in$(pos C)$\cap$(pos D). We will choose points from both C and D to form a minimal positive basis for a minimal subspace, thereby denying d and completing the proof. Let C^+ denote the subset of C which appears in the unique minimal representation of $x\in$pos C. Let D^- denote the subset of D which appears in the unique minimal representation of $-x\in$pos D. Let B_1 denote $C^+\cup D^-$. B_1 is positively independent because it is a subset of B. It is easily seen that $C^+\cup\{-x\}$ and $D^-\cup\{x\}$ are minimal positive bases for the linear subspaces pos $(C^+\cup\{-x\})$ and pos $(D^-\cup\{x\})$, respectively, of pos B_1. Since B_1 is contained in and positively spans the linear sum of these two subspaces, it follows that pos B_1 is a linear space, with B_1 as a positive basis. To show that B_1 is a minimal positive basis, it suffices to show that card $B_1=$(dim pos B_1)$+1$. The denial would assert that card $B_1>$dim pos B_1+1. But then if we choose a point $c\in C^+\subset B_1$, there exists a subset $M\subset B_1$, such that $c\in M$ and M is a minimal positive basis for the minimal space pos M. Now since M is a positive

basis, and therefore admits positive relations, it follows that M can not be a subset of C^+. Thus M has elements from both C^+ and D^-. This completes the proof of the lemma.

The following example shows that if dim $L=n\geqq 5$, then condition g does not imply the uniqueness of minimal representations, and therefore Lemma 3 cannot be extended to higher dimensional spaces. The positive basis $B=\{b_i \mid 1\leqq i\leqq 8\}$ has 4 subsets which are minimal positive bases for minimal subspaces, namely

$$B_1 = \{b_1, b_2, b_5, b_6\}, \quad B_2 = \{b_3, b_4, b_7, b_8\}, \quad B_3 = \{b_1, b_2, b_7, b_8\} \text{ and } B_4 = \{b_3, b_4, b_5, b_6\}.$$

Also, $B=B_1\cup B_2$ and $B=B_3\cup B_4$ are two partitionings of B into minimal positive bases. But the point (1, 0, 0, 0, 0) does not have a unique minimal representation.

$$\begin{aligned} b_1 &= (1, 1, 0, 0, 0) & b_5 &= (-1, 0, 0, 1, 0)\\ b_2 &= (1, -1, 0, 0, 0) & b_6 &= (-1, 0, 0, -1, 0)\\ b_3 &= (1, 0, 1, 0, 0) & b_7 &= (-1, 0, 0, 0, 1)\\ b_4 &= (1, 0, -1, 0, 0) & b_8 &= (-1, 0, 0, 0, -1). \end{aligned}$$

The following example shows that if dim $L=n\geqq 6$, then condition f does not imply the uniqueness of minimal positive representations, and therefore Lemma 4 cannot be extended to higher dimensional spaces. The positive basis $B=\{b_i \mid 1\leqq i\leqq 9\}$ has 3 subsets which are minimal positive bases for minimal subspaces, namely

$$B_1 = \{b_1, b_2, b_3, b_4\}, \quad B_2 = \{b_5, b_6, b_7, b_8, b_9\} \quad \text{and} \quad B_3 = \{b_1, b_2, b_5, b_6, b_8\}.$$

Also, $B=B_1\cup B_2$ is a unique partition of B into minimal subspaces. But each minimal representation of the point (0, 0, 0, -4, -4, 0) has length 3, and there are 2 distinct such representations.

$$\begin{aligned} b_1 &= (1, 0, 0, 0, 0, 0) & b_6 &= (0, 0, 0, 0, 1, 0)\\ b_2 &= (0, 1, 0, 0, 0, 0) & b_7 &= (0, 0, 0, 0, 0, 1)\\ b_3 &= (0, 0, 1, 0, 0, 0) & b_8 &= (-1, -1, 0, -4, -4, 0)\\ b_4 &= (-1, -1, -1, 0, 0, 0) & b_9 &= (1, 1, 0, 2, 2, -2).\\ b_5 &= (0, 0, 0, 1, 0, 0) \end{aligned}$$

3. Continuous representations. If a minimal representation $\alpha^x\in\mathcal{P}B$ of each $x\in L$ is unique, and L is of finite dimension n, then standard arguments show that the map $x\to\alpha^x$ (from n to (card B)-dimensional euclidean space) is bicontinuous. This is analogous, of course, to the situation concerning the unique representations of points in terms of an arbitrary linear basis. On the other

hand, if minimal representations are not unique, then it is easy to show that it is never possible to assign a minimal representation to each point of the space in such a way that the representation varies continuously with the point. For example, if $\{\delta_i\}$ is a sequence of positive numbers approaching zero, then the sequences $\{(1, \delta_i, 0, 0, 0)\}$ and $\{(1, 0, \delta_i, 0, 0)\}$ from the first example of the previous section both converge to the point (1, 0, 0, 0, 0). And yet, each point in each sequence has a unique minimal representation involving only b_1 and b_2 for the first sequence, and involving only b_3 and b_4 for the second sequence.

Even if minimal representations are not unique, there is a rather natural technique that assigns a well-defined representation (in terms of any given positive basis) to each point of the space in such a way that the representation varies continuously with the point. If minimal representations are unique, the assigned representation will always be the unique minimal representation. The following lemma, which provides the necessary machinery, has also been used to prove certain Helly-type convexity theorems, (see [4]). Note that if L is the direct linear sum of its minimal subspaces, then the partitioning of B produced by the lemma is unique, and is precisely the partition guaranteed by condition f of the last section.

LEMMA 6. *Let B be any positive basis for (finite-dimensional) L. Then B admits a partition into pair-wise disjoint subsets $B=B_1\cup \cdots \cup B_k$ such that card $B_i \geq$ card $B_{i+i} \geq 2$ and pos $(B_1\cup \cdots \cup B_j)$ is a linear subspace of L of dimension $(\sum_{i=1}^{j}$ card $B_i)-j$ for each $j=1, 2, \cdots, k$.*

Sketch of Proof. Let B_1 be a subset of B of maximal cardinality such that B_1 is a minimal positive basis for the minimal subspace pos B_1. If pos $B_1 \neq L$, let L_1 be a supplementary linear subspace, so that $L=L_1\oplus$pos B_1, and let π_1 denote the natural projection of L onto L_1. It follows that $\pi_1(B\sim B_1)$ is a positive basis for L_1. Also, if B_2 is a subset of B of maximal cardinality such that $\pi_1 B_2$ is a minimal positive basis for the minimal subspace pos $\pi_1 B_2$ of L_1, then it follows that card $B_1 \geq$ card B_2 and pos $(B_1\cup\pi_1 B_2)=$pos $(B_1\cup B_2)$. (See [4] for detailed arguments.) It is clear that the conditions of the lemma are satisfied. If $B\neq B_1\cup B_2$ we continue with the same argument: choose L_2 such that $L_1=L_2 \oplus$pos $\pi_1 B_2$, let $\pi_2: L_1\to L_2$ be the natural projection, and let B_3 be the largest subset of $B\sim(B_1\cup B_2)$ such that $\pi_2\pi_1 B_3$ is a minimal positive basis for the minimal subspace pos $\pi_2\pi_1 B_3$ of L_2. This process is continued until the Lemma is established.

In the above notation it is clear that

$$P = B_1 \cup (\pi_1 B_2) \cup (\pi_2\pi_1 B_3) \cup \cdots \cup (\pi_{k-1} \cdots \pi_1 B_k)$$

is a positive basis for L and each minimal subspace is determined by one of the sets $(\pi_{j-1} \cdots \pi_1 B_j)$. Since by construction these minimal subspaces have only $\{0\}$ in common, it follows by Theorem 1 that a minimal representation α of a point $x\in L$ in terms of P is unique. It is this uniquely defined minimal representation α of the point $x\in L$ (in terms of P) that we use to obtain a well-defined

representation for x in terms of B. Let $P(i)$ denote the set $(\pi_{i-1} \cdots \pi_1 B_i)$ for $i=1, \cdots, k$. Then

$$x = \sum_{i=1}^{k}\left(\sum_{b\in P(i)} \alpha_b b\right) \quad \text{and} \quad \sum_{b\in P(k)} \alpha_b b = e_k + \sum_{b\in P(k)} \alpha_b(\pi_{k-1})^{-1}b,$$

where $e_k \in \text{pos } P(k-1)$. We combine e_k with $(\sum_{b\in P(k-1)} \alpha_b b) \in \text{pos } P(k-1)$ to obtain a point $e_{k-1} \in \text{pos } P(k-1)$. Since by definition,

$$\begin{aligned} L_{k-2} &= \text{pos } P(k-1) \oplus \text{pos } P(k) = \text{pos }(P(k-1) \cup P(k)) \\ &= \text{pos }(P(k-1) \cup (\pi_{k-1})^{-1}P(k)), \end{aligned}$$

we note that $(\pi_{k-1})^{-1}P(k) \cup P(k-1)$ is exactly $(\pi_{k-2} \cdots \pi_1(B \sim \bigcup_{j=1}^{k-2} B_j))$, the original positive basis of L_{k-2}. Since pos $P(k-1)$ is a minimal subspace of L_{k-2}, it follows that $e_{k-1} \in \text{pos } P(k-1)$ has a unique minimal representation $e_{k-1} = \sum_{b\in P(k-1)} \beta_b b$, and

$$\begin{aligned} \left(\sum_{b\in P(k-1)} \alpha_b b\right) + \left(\sum_{b\in P(k)} \alpha_b b\right) &= e_{k-1} + \left(\sum_{b\in P(k-1)} \alpha_b(\pi_k)^{-1}b\right) \\ &= \left(\sum_{b\in P(k-1)} \beta_b b\right) + \left(\sum_{b\in(\pi_k^{-1}P(k))} \alpha_b b\right). \qquad (*) \end{aligned}$$

Since α varies continuously with x, it follows that e_k, e_{k-1} and therefore also the coefficients of the last equation vary continuously with x. The representation (*) is an element of L_{k-2}, and is clearly a well-defined positive representation in terms of the positive basis $P(k-1) \cup (\pi_{k-1})^{-1}P(k)$ of L_{k-2}. This argument may now be repeated for the space $L_{k-2} \oplus \text{pos } P(k-2)$ and so on, until the desired continuous representation of x is obtained.

If $B = B_1 \cup \cdots \cup B_k$ is a partition of B as in Lemma 6, and if for some j, $1 \leqq j \leqq k$, pos $(B_1 \cup \cdots \cup B_j)$ the direct linear sum of the minimal subspaces pos B_i, $1 \leqq i \leqq j$, then the space L_i may be chosen to include the sets $B_i, 1 < i \leqq j$. It then follows that the continuous representation α chosen above for a point $x \in \text{pos } (B_1 \cup \cdots \cup B_j)$ is precisely the (unique) minimal representation of x in terms of B.

This research was supported in part by the N.S.F. Program for the College-University Association for Improvement in the Teaching of Science in the Pacific Northwest. The author is much indebted to Professors Micha Perles and R. L. McKinney for their valuable correspondences on the topics of this paper. The second example of section 2 is due to McKinney.

References

1. W. Bonnice and V. L. Klee, The generation of convex hulls, Math. Ann., 152 (1963) 1–29.

2. Chandler Davis, Theory of positive linear dependence, Amer. J. Math., 76 (1954) 733–746.

3. R. L. McKinney, Positive bases for linear spaces, Trans. Amer. Math. Soc., 103 (1962) 131–148.

4. J. R. Reay, Generalizations of a theorem of Caratheodory, Amer. Math. Soc. Memoir, No. 54, 1965.

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

THE HAT-CHECK PROBLEM

RICHARD SCOVILLE, Duke University

The hat-check problem is a familiar one in elementary probability theory: a hat-check girl in a restaurant, having checked n hats, gets them hopelessly scrambled and returns them at random to the n owners as they leave. What is the probability p_n that nobody gets his own hat back? The question can be posed as a card problem and as such was treated by many of the early writers in probability under the name *Treize* or *Rencontre* [1]: a deck of cards numbered from 1 to n is shuffled and dealt. What is the probability p_n that no card is in its correct place, that is, that card i is not ith from the top for any i? The answer is $p_n=1/2!-1/3!+1/4!- \cdots \pm 1/n!$ so that in particular $\lim p_n=1/e$.

Let us turn to the more practical problem of considering a deck of kn cards consisting of k suits of n denominations (1 to n) each. We shuffle this deck, deal the cards and at the same time count from 1 to n, k times. What is the probability $p_{n,k}$ that no card matches the denomination called? In this case the shuffle (or permutation) of the deck will be called *complete*. The number $p_{n,k}$ is rather complicated but we are at least able to prove the following

THEOREM 1. *For fixed k, $\lim_n p_{n,k}=1/e^k$.*

Hence the answer is essentially the same for large n as the answer we would obtain if the deck were first divided into suits and each suit were dealt separately. It is also interesting that if we replace each card in the deck after dealing it the probability is $(1-1/n)^{kn}$ so in this case, the independent case, we also get $1/e^k$ as the limiting value.

Let k and n be fixed for a while. Consider the $k\times k$ matrix $S=(s_{ij})$ with non-negative integral entries. Let $B(S)$ be the number of ways of splitting the deck into k disjoint sets $A_1, A_2, \cdots, A_k$ each containing n cards, such that there are s_{ij} i-of-a-kinds in part A_j, that is, such that there are s_{ij} denominations represented in A_j by precisely i suits. It is clear that

$$(kn)! = \sum_S B(S)(n!)^k, \tag{1}$$

since every possible permutation of the deck is determined by the *sets* A_1 (the first n cards), A_2 (the second n cards), etc.; and a permutation of each of these sets. If $k=2$ this is the well-known formula

$$(2n)! = \sum_{j=0}^{[n/2]} \binom{n}{j}\binom{n-j}{j} 2^{n-2j}(n!)^2$$

because in this case if $B(S)\neq 0$ then S is of the form

$$S = \begin{bmatrix} n - 2j & n - 2j \\ j & j \end{bmatrix}$$

for some j between 0 and $[n/2]$, and

$$B(S) = \binom{n}{j}\binom{n-j}{j}2^{n-2j},$$

the number of possible choices of suits and denominations. Now of the $(n!)^k$ permutations of the deck in which the first n cards are the cards of the set A_1, the next n cards are the cards of the set A_2, etc., only $M(A_1)\cdot M(A_2) \cdots M(A_k)$ are complete permutations, where $M(A_j)$ is the number of permutations of the set A_j which are complete, i.e., in which the ith card is not an i for any i. $M(A_j)$, however, depends only on the numbers $s_{1j}, s_{2j}, \cdots, s_{kj}$ and not on the particular suits and denominations appearing in A_j. Because of this we can define

$$M(s_{1j}, s_{2j}, \cdots, s_{kj}) = M(A_j).$$

The number of complete permutations of the deck is therefore given by

$$\sum_S B(S) \prod_{j=1}^{k} M(s_{1j}, s_{2j}, \cdots, s_{kj})$$

and

$$p_{n,k} = \frac{(n!)^k}{(kn)!} \sum_S B(S) \prod_{j=1}^{k} (M(s_{1j}, \cdots, s_{kj})/n!)$$

which is seen from (1) to be a convex combination of the numbers

$$\prod_{j=1}^{k} (M(s_{1j}, s_{2j}, \cdots, s_{kj})/n!).$$

Hence our theorem will follow immediately from the following

THEOREM 2. *For every* $k \geqq 1$, *we have* $\lim_n M(s_1, \cdots, s_k)/n! = 1/e$ *uniformly for all sequences* $\{s_i\}$ *of nonnegative integers satisfying*

$$\sum_{i=1}^{k} is_i = n.$$

Proof. Notice first that the number $M(s_1, s_2, \cdots, s_k)$ is the same as the number $M(s_1, s_2, \cdots, s_{k-1})$ provided $s_k = 0$.

The proof of the theorem will proceed by induction on k. For $k=1$ only the value $s_1 = n$ is possible for each n and the conclusion of the theorem is simply the result stated earlier. Suppose that the theorem is true for all numbers less than k. We will prove that if $s_k \neq 0$ and $k > 2$ then

(2) $0 \leqq M(s_1 + 1, s_2, \cdots, s_{k-1} + 1, s_k - 1) - M(s_1, \cdots, s_k) \leqq (k-1)(n-2)!$

while if $k=2$ then

$$(2') \qquad 0 \leqq M(s_1+2, s_2-1) - M(s_1, s_2) \leqq (n-2)!,$$

both formulas holding for all $n>2$. Let $C(i, j)$ be the card from suit i having denomination j. Consider the collection A of n cards defined by

$$A = \{C(i,j) \mid 1 \leqq j \leqq s_k + \cdots + s_{k-i+1},\ 1 \leqq i \leqq k\}.$$

We clearly have $M(A)=M(s_1, s_2, \cdots, s_k)$ where, as before, $M(A)$ is the number of complete permutations of A. Consider also the collection B of n cards consisting of all the cards of A except that $C(1, 1)$ is removed and replaced by $C(k, s_k+\cdots+s_1+1)$. For convenience let $u=s_k+\cdots+s_1+1$. Note that $M(B)=M(s_1+1, \cdots, s_{k-1}+1, s_k-1)$ if $k>2$ and $M(B)=M(s_1+2, s_2-1)$ if $k=2$.

The point of this consideration is that each complete permutation of A may be turned into a different complete permutation of B and the complete permutations of B not obtained in this way are fewer than $(k-1)(n-2)!$ in number: Let $\pi=(\pi_1, \pi_2, \cdots, \pi_n)$ be a complete permutation of A. Suppose first that $C(1, 1)$ is not in the uth place in π. Then it is only necessary to replace $C(1, 1)$ by $C(k, u)$ to get a complete permutation of B. The other possibility is that $C(1, 1)$ is in the uth place in π. In this case put $C(k, u)$ in the first place and put π_1 in the uth place. Hence we obtain again a complete permutation of B. This shows that $M(A)\leqq M(B)$. What are the complete permutations of B that are not obtained in this way? They are those with $C(k, u)$ in the first place and in the uth place one of the $k-1$ cards $C(2, 1), C(3, 1), \cdots, C(k, 1)$ of denomination 1 in B. Obviously there are fewer than $(k-1)(n-2)!$ of these. This proves (2) and (2′).

Now, given $\epsilon>0$, use the inductive hypothesis to choose $n_0>2/\epsilon+1$ such that

$$|M(t_1, \cdots, t_{k-1})/n! - 1/e| < \epsilon/2$$

for any $n>n_0$ and any sequence $\{t_i\}$ satisfying $\sum_{i=1}^{k-1} it_i=n$. For any sequence $\{s_i\}$ satisfying $\sum_{i=1}^{k} is_i=n$ we have, for $k>2$,

$$\begin{aligned} &M(s_1+s_k, s_2, \cdots, s_{k-1}+s_k, 0) - s_k(k-1)(n-2)! \\ &\qquad \leqq M(s_1, \cdots, s_k) \leqq M(s_1+s_k, s_2, \cdots, s_{k-1}+s_k, 0) \end{aligned}$$

by applying (2) repeatedly, so that

$$\frac{M(s_1+s_k, \cdots, s_{k-1}+s_k, 0)}{n!} - \frac{1}{n-1} \leqq \frac{M(s_1, \cdots, s_k)}{n!} \leqq \frac{M(s_1+s_k, \cdots, s_{k-1}+s_k, 0)}{n!}$$

since $s_k \leqq n/k$. Hence,

$$|M(s_1, \cdots, s_k)/n! - 1/e| < \epsilon.$$

This proves the theorem.

This work was supported by a contract from the National Science Foundation.

Reference

1. I. Todhunter, A History of the Mathematical Theory of Probability, Chelsea reprint, New York, 1949.

A STABILITY PROBLEM IN NUMBER THEORY

J. D. DIXON, California Institute of Technology, Pasadena
(Presently at the University of New South Wales, Australia)

In his book [2, Chapter VI], S. M. Ulam discusses problems which arise as natural analogues of questions of stability of solutions in analytic problems. The two theorems presented here deal with such a problem arising in number theory.

As motivation, consider first the following simple example. Let g_1 and g_2 be two functions which, like all functions considered in this paper, are defined from the natural numbers into the complex numbers. Suppose f_1 and f_2 are the respective solutions of the infinite sets of equations

$$g_i(n) = \sum_{d|n} f_i(d) \qquad (n = 1, 2 \cdots) \quad \text{for } i = 1, 2.$$

Then the Möbius inversion formula [1, Section 16.4] shows that

$$f_i(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g_i(d) \qquad (n = 1, 2 \cdots) \quad \text{for } i = 1, 2.$$

Now suppose that $g_2(n) = g_1(n) + o(1)$ as $n \to \infty$. Then (cf. the proof of Theorem 2), we find that $f_2(n) = f_1(n) + o(\tau(n))$ as $\tau(n) \to \infty$, where $\tau(n)$ is the number of divisors of n. In general, the error term cannot be improved very much. Specifically, Theorem 2 will show that we can have $f_2(n) = f_1(n) + o(1)$ only if $f_1 = f_2$ and $g_1 = g_2$. Alternatively, if $g_1 \neq g_2$ then

$$\limsup_{n \to \infty} |f(n) - f_1(n)| > 0.$$

As a second example of Theorem 2 we have the following result: For real $\alpha > 0$, we write $\sigma_\alpha(n) = \sum_{d|n} d^\alpha$. Then the only function f such that

$$\text{(i)} \qquad \sum_{d|n} \sigma_\alpha(d) f\left(\frac{n}{d}\right) \sim n^\alpha, \qquad \text{(ii)} \quad f(n)/n^\alpha \to 0,$$

as $n \to \infty$, is the Möbius function μ.

Notation. A sequence $(a_n)_{n \geq 1}$ is called *multiplicative* if not all its terms are zero, and $a_{mn} = a_m a_n$ whenever the greatest common divisor $(m, n) = 1$. In particular, $a_1 = 1$.

We shall use $\tau(n)$ and $\nu(n)$ to denote, respectively, the number of divisors of n, and the number of prime factors (counting multiplicities) of n. As usual, μ is the Möbius function and n is called *squarefree* if $\mu(n) \neq 0$.

We shall prove the following theorems:

THEOREM 1. *Let $(a_n)_{n\geq 1}$ be a multiplicative sequence of complex numbers such that for some infinite sequence P of distinct prime numbers, $a_p \to 1$ as p runs through P. Let B be any function such that $B(m)\to 0$ as $m\to\infty$. If a function f satisfies the conditions*

$$\sum_{d|n} a_d f\left(\frac{n}{d}\right) \to 0 \qquad \text{as } n\to\infty; \tag{1}$$

$$\left|\frac{1}{\tau(n)} \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right)\right| \leqq B(\tau(n)) \qquad \text{for all } n; \tag{2}$$

then f is identically zero.

THEOREM 2. *Let $(b_n)_{n\geq 1}$ be a multiplicative sequence of complex numbers such that for some infinite sequence P of distinct prime numbers, and some real α, $b_p/p^\alpha \to 1$ as p runs through P.*

If f, f_1 and g are functions which satisfy the conditions

$$\sum_{d|n} b_d f\left(\frac{n}{d}\right) = g(n) + o(n^\alpha) \qquad \text{as } n\to\infty; \tag{3}$$

$$\sum_{d|n} b_d f_1\left(\frac{n}{d}\right) = g(n) \qquad \text{for all } n; \tag{4}$$

$$f(n) = f_1(n) + o(n^\alpha) \qquad \text{as } n\to\infty; \tag{5}$$

then $f=f_1$.

Proof of Theorem 1. We have to show that $f(n)=0$ for each n.

For each pair of positive integers n and ν, we let $S_n(\nu)$ denote any sequence $(m_k)_{k\geq 1}$ with the properties:

(i) each m_k is a positive integer relatively prime to n;

(ii) each m_k is squarefree and has exactly ν prime factors, each of which appears in the sequence P;

(iii) as $k\to\infty$, all the prime factors of m_k tend to infinity.

The proof of the theorem now proceeds in several steps.

(A) *As m runs through $S_n(\nu)$* and $n\to\infty$, $\sum_{d|m} a_d f(mn/d)\to 0$.

We proceed by induction on $\nu(n)$. For $\nu(n)=0$, the result follows from (1). In general, for $\nu(n)>0$,

$$\sum_{d|m} a_d f\left(\frac{mn}{d}\right) = \sum_{d|mn} a_d f\left(\frac{mn}{d}\right) - \sum_{c|n, c\neq 1} a_c \sum_{d|m} a_d f\left(\frac{mn}{dc}\right),$$

because $(a_n)_n$ is a multiplicative sequence and $(c, d)=1$ for $c|n$ and $d|m$. By applying the induction hypothesis to n/c (for $c|n$, $c\neq 1$) we see that the last sum tends to zero, while the first sum on the right hand side tends to zero by (1).

(B) *As m runs through $S_n(\nu)$ and $n\to\infty$, $f(mn)\to\mu(m)f(n)$.*

(Note: $\mu(m)=(-1)^\nu$ because m is squarefree.)

We proceed by induction on ν. $f(n)=\mu(1)f(n)$ provides the first step and in general, for $\nu>0$,

$$f(mn)-\mu(m)f(n) = \sum_{d|m} a_d f\left(\frac{mn}{d}\right) - \sum_{d|m,d\neq 1} a_d\left\{f\left(\frac{mn}{d}\right)-\mu\left(\frac{m}{d}\right)f(n)\right\}$$
$$-\sum_{d|m}\mu\left(\frac{m}{d}\right)a_d f(n).$$

(Note that, because $(a_n)_n$ is multiplicative, $a_1=1$.) On the right side, the first sum goes to zero by (A), and the second sum goes to zero by the induction hypothesis applied to m/d (for $d|m$, $d\neq 1$). Because the sequence $(a_n)_n$ is multiplicative and m is squarefree, the last term on the right hand side equals $\mu(m)f(n)\Pi_{p|m}(1-a_p)$ (where p runs through the prime divisors of m). By the initial hypothesis on the sequence $(a_n)_n$ and the definition of $S_n(\nu)$, it follows that $a_p\to 1$ for each p dividing m as m runs through $S_n(\nu)$. Thus the last term of the right hand side also tends to zero and the assertion follows.

(C) *For each positive integer n, there exists a function B_n such that $B_n(\nu)\to 0$ as $\nu\to\infty$, and*

$$\left|\frac{1}{\tau(m)}\sum_{d|m}\mu(d)f\left(\frac{mn}{d}\right)\right| \leqq B_n(\nu) \qquad \textit{for all } m \textit{ in } S_n(\nu). \tag{6}$$

We proceed by induction on $\nu(n)$. For $\nu(n)=0$, it follows from (2) that we may take $B_1(\nu)=B(\tau(m))$. ($\tau(m)=2^\nu$ because m is square-free.) In general, for $\nu(n)>0$,

$$\sum_{d|m}\mu(d)f\left(\frac{mn}{d}\right) = \sum_{d|mn}\mu(d)f\left(\frac{mn}{d}\right) - \sum_{c|n,c\neq 1}\mu(c)\sum_{d|m}\mu(d)f\left(\frac{mn}{d}\right),$$

because μ is a multiplicative function and $(c,d)=1$ for all $c|n$ and $d|m$. Using (2) and the induction hypothesis applied to n/c (for $c|n$, $c\neq 1$), we obtain (6) with

$$B_n(\nu) = B(2^\nu\tau(n))\tau(n) + \sum_{c|n,c\neq n} B_c(\nu) \to 0 \qquad \text{as } \nu\to\infty.$$

(D) Finally, we show that $f(n)=0$.

For squarefree m,

$$\sum_{d|m}\mu(d)\mu\left(\frac{m}{d}\right) = \mu(m)\tau(m).$$

Thus, as m runs through $S_n(\nu)$, we have

$$\mu(m)\tau(m)f(n) = \sum_{d|m}\mu(d)f\left(\frac{mn}{d}\right) - \sum_{d|m}\mu(d)\left\{f\left(\frac{mn}{d}\right)-\mu\left(\frac{m}{d}\right)f(n)\right\}.$$

From (B), it follows that the second sum on the right hand side tends to zero. Therefore, by (C), $|f(n)| \leqq B_n(\nu)$. But $B_n(\nu)\to 0$ as $\nu\to\infty$, so $f(n)=0$. Thus Theorem 1 is proved.

Proof of Theorem 2. For all n we define

$$a_n = n^{-\alpha}b_n \quad \text{and} \quad h(n) = n^{-\alpha}\{f(n) - f_1(n)\}.$$

Then the sequence $(a_n)_n$ satisfies the hypotheses of Theorem 1, and

$$\sum_{d|n} a_d h\left(\frac{n}{d}\right) = n^{-\alpha}\sum_{d|n} b_d\left\{f\left(\frac{n}{d}\right) - f_1\left(\frac{n}{d}\right)\right\} \to 0 \qquad \text{as } n\to\infty.$$

Also, for each $\epsilon>0$, there exists an integer $N>0$ such that $|h(n)| = n^{-\alpha}|f(n) - f_1(n)| < \epsilon/2$ for $n>N$. Hence

$$\left|\sum_{d|n}\mu(d)h\left(\frac{n}{d}\right)\right| \leqq \sum_{d|n} |h(d)| < \tau(n)\epsilon/2 + \sum_{d=1}^{N} |h(d)| < \epsilon\tau(n),$$

when $\tau(n)$ is sufficiently large. Thus we can find a suitable function B in order to be able to apply Theorem 1 to show that h is identically zero. Then $f=f_1$, and Theorem 2 is proved.

Note. In case $a_n=1$ for all n, the following examples serve to show that conditions (1) and (2) of Theorem 1 cannot be relaxed.

(a) If it is only supposed that B is bounded (rather than $o(1)$), then the function $f=\mu$ satisfies the two conditions, with $B(m)=1$.

(b) If it is only supposed that the sums in (1) are bounded (rather than $o(1)$), then the function f, where $f(n)=n^{-2}$ for all n, satisfies this condition and condition (2).

References

1. G. H. Hardy and E. M. Wright, Theory of Numbers, Clarendon Press, 1954.
2. S. M. Ulam, Collection of Mathematical Problems, Interscience. New York. 1960.

ON CERTAIN INEQUALITIES FOR THE CHARACTERISTIC ROOTS OF HERMITIAN MATRICES

K. N. MAJINDAR, Loyola College, Montreal

All the matrices here have their elements in the field of complex numbers. For any matrix P, P^* denotes its conjugate transpose, and likewise for vectors. If H_1 and H_2 are two nth order hermitian matrices then we write $H_1\leqq H_2$ if and only if $x^*H_1x\leqq x^*H_2x$ for all column vectors x with n components. Recently Everitt [1] has given new proofs for the following theorems.

THEOREM 1. *Let H_1 and H_2 be two n-th order matrices with characteristic roots $\lambda_1\leqq\lambda_2\leqq\lambda_3\leqq\cdots\leqq\lambda_n$ and $\Lambda_1\leqq\Lambda_2\leqq\cdots\leqq\Lambda_n$ respectively. If $H_1\leqq H_2$ then $\lambda_r\leqq\Lambda_r$, $r=1, 2, \cdots, n$.*

THEOREM 2. *Let H be a hermitian matrix of order m, and characteristic roots $\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_m$. If H_1 is any principal $(m-1)\times(m-1)$ submatrix of H, with characteristic roots $\alpha_1' \leqq \alpha_2' \leqq \cdots \leqq \alpha_{m-1}'$, then $\alpha_r \leqq \alpha_r' \leqq \alpha_{r+1}$, $r=1, 2, \cdots, (m-1)$.*

Everitt has shown that either of these theorems can be deduced from the other. We will give here a theorem from which both these theorems can be deduced as easy corollaries. Our theorem can be stated as follows.

THEOREM. *Let A and B be hermitian matrices of orders n and N respectively and with respective characteristic roots $\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_n$ and $\Lambda_1 \leqq \Lambda_2 \leqq \cdots \leqq \Lambda_N$, $n \leqq N$. If $A \leqq C^* BC$ then $\lambda_r \leqq \Lambda_{N-n+r}$, and if $A \geqq C^* BC$ then $\lambda_r \geqq \Lambda_r$, $r=1, 2, \cdots, n$ where C is an $N\times n$ matrix such that $C^* C=I_n$, (I_n denoting the identity matrix of order n).*

Proof. Let U and V be unitary matrices of order n and N respectively which transform A and B to the diagonal matrices with diagonal elements $\lambda_1, \lambda_2, \cdots, \lambda_n$ and $\Lambda_1, \Lambda_2, \cdots, \Lambda_N$ coming in this order. Thus $U^* AU=D_1$ and $V^* BV=D_2$.

Suppose first that $A \leqq C^* BC$. Now $A\leqq C^* BC$ if and only if $D_1 \leqq E^* D_2E$ where $E=VCU$. Clearly $E^* E=I_n$. Let $E=(e_{ij})$, $i=1, 2, \cdots, N, j=1, 2, \cdots, n$. If x is any column vector with elements $x_1, x_2, \cdots, x_n$ then we have $x^* D_1x \leqq z^* D_2z$ where $z=Ex$. If the elements of z are $z_1, z_2, \cdots, z_N$, then

$$z_i = e_{i1}x_1 + e_{i2}x_2 + \cdots + e_{in}x_n, \qquad i = 1, 2, \cdots, N.$$

Further we have $\sum_1^N z_i\bar{z}_i=z^* z=x^* E^* Ex=x^* x=\sum_1^n x_i\bar{x}_i$.

From the relation $x^* D_1x\leqq z^* D_2z$ we get $\sum_1^n \lambda_i x_i\bar{x}_i \leqq \sum_1^N z_i\bar{z}_i\Lambda_i$. Choose now $x_1=x_2=\cdots=x_{r-1}=0$. Determine the rest of the x_i's nontrivially in such a way that

$$e_{ir}x_r + e_{ir+1}x_{r+1} + \cdots + e_{in}x_n = 0,\ i = N-n+r+1, N-n+r+2, \cdots, N.$$

Such a determination is possible as we have $n-r$ linear homogeneous equations in $n-r+1$ unknowns. Let these values of the x_i's be $x_1', x_2', \cdots, x_n'$ and the corresponding values of the z_i's be $z_1', z_2', \cdots, z_N'$. Then

$$z'_{N-n+r+1} = z'_{N-n+r+2} = \cdots = z'_N = 0.$$

Thus we obtain

$$\sum_r^n \lambda_i x_i'\bar{x}_i' \leqq \sum_1^{N-n+r} \Lambda_i z_i'\bar{z}_i', \quad \text{i.e., } \lambda_r \sum_1^n x_i'\bar{x}_i' \leqq \Lambda_{N-n+r} \sum_1^N z_i'\bar{z}_i',$$

or $\lambda_r \leqq \Lambda_{N-n+r}$, $r=1, 2, \cdots, n$.

Suppose next that $A \geqq C^* BC$. Proceeding as above and choosing $x_{r+1} = \cdots = x_n=0$ and the rest of the x_i's in such a way that the corresponding z_i, $i=1, 2, \cdots, r-1$ become 0 we get similarly $\lambda_r \geqq \Lambda_r$, $r=1, 2, \cdots, n$. This completes the proof.

To deduce Theorem 1, take $n=N$, $A=H_1$, $B=H_2$, $C=I_n$ in our theorem. For Theorem 2, let us assume without loss of generality that H_1 is the principal submatrix obtained by suppressing the first row and column of H. Let C be the $m\times(m-1)$ matrix got by augmenting the matrix I_{m-1} by a top row of $m-1$ zeros. Then $H_1=C^*BC$. So by taking $N=m$, $n=m-1$ in our theorem, we get $\alpha_r' \leqq \alpha_{r+1}$ and $\alpha_r' \geqq \alpha_r$, $r=1, 2, \cdots, (m-1)$.

Reference

1. W. N. Everitt, Two theorems in matrix theory, this MONTHLY, 69 (1962) 856–859.

AN IDENTITY FOR THIRD ORDER DETERMINANTS

HARLEY FLANDERS, Purdue University

Suppose that $A=\|a_{ij}\|$ is a three by three non-singular matrix and that $C=A^{-1}=\|c_{ij}\|$. Assume that each a_{ij} and each c_{ij} is not zero. We assert

THEOREM 1. *The matrices of reciprocals* $\|a_{ij}^{-1}\|$ *and* $\|c_{ij}^{-1}\|$ *are singular or non-singular together.*

This has a geometric interpretation as follows.

Let A_1, A_2, A_3 *be three points of the projective plane and* K_1, K_2, K_3 *three (non-degenerate) conics, each containing the three given points. The conics* K_2 *and* K_3 *intersect in a fourth point* Q_1. *Determine* Q_2, Q_3 *similarly and suppose* Q_1, Q_2, Q_3 *to be distinct (i.e., the conics are supposed linearly independent). Let* P_1, P_2, P_3 *be the poles of* K_1, K_2, K_3 *respectively with respect to the line* $\overline{A_1A_2}$. *Conclusion: The points* Q_1, Q_2, Q_3 *are collinear if and only if the six points* A_1, A_2, A_3, P_1, P_2, P_3 *lie on a conic.*

This follows from Theorem 1 by choosing homogeneous coordinates conveniently. Take $A_1=(1, 0, 0)$ $A_2=(0, 1, 0)$, $A_3(0, 0, 1)$. Then the equation of K_i is

$$a_{i1}x_2x_3 + a_{i2}x_3x_1 + a_{i3}x_1x_2 = 0,$$

and we have our matrix $A=\|a_{ij}\|$. It is nonsingular because the conics are independent; no a_{ij} vanishes or else K_i is degenerate. The other intersection Q_1 of K_2 and K_3 is found, after a short computation, to be

$$Q_1 = (c_{11}^{-1}, c_{21}^{-1}, c_{31}^{-1}).$$

[Actually one obtains $Q_1=(c_{21}c_{31}, c_{31}c_{11}, c_{11}c_{21})$. No coordinate can vanish since a line like $x_1=0$ cannot have three points in common with a K.] It follows that the condition for collinearity of Q_1, Q_2, Q_3 is det $\|1/c_{ij}\|=0$.

On the other hand, the pole P_i of $x_3=0$ with respect to K_i is

$$P_i=(a_{i1}, a_{i2}, -a_{i3}).$$

The most general conic containing A_1, A_2, A_3 is

$$b_1x_2x_3 + b_2x_3x_1 + b_3x_1x_2 = 0,$$

where $(b_1, b_2, b_3) \neq (0, 0, 0)$. We want such a conic to contain P_1, P_2, and P_3. Substitution of the coordinates of the P_i leads to a system of three homogeneous linear equations for the b_i. Necessary and sufficient for a nontrivial solution is precisely the condition

$$\det \|1/a_{ij}\| = 0.$$

[N.B. The problem on circles of F. J. Servedio, (this MONTHLY, May 1964, p. 554) is obtained when one specializes the result above to the case in which A_1, A_2 are the circular points at infinity.]

We shall reformulate Theorem 1 as a polynomial identity not involving reciprocals. We recall that if A is an n by n matrix, we can define the cofactor matrix (sometimes called adjoint) cof A. It has these properties:

$$\begin{aligned} A \operatorname{cof} A &= (\det A)I. \\ \det(\operatorname{cof} A) &= (\det A)^{n-1} \\ \operatorname{cof}(\operatorname{cof} A) &= (\det A)^{n-2}A. \\ (\operatorname{cof} A)/(\det A) &= A^{-1} \text{ if } A \text{ is nonsingular.} \end{aligned}$$

Let $A = \|a_{ij}\|$ be 3×3 and $B = \operatorname{cof} A = \|b_{ij}\|$. Set

$$A_r = \left\| \begin{matrix} a_{12}a_{13} & a_{13}a_{11} & a_{11}a_{12} \\ a_{22}a_{23} & a_{23}a_{21} & a_{21}a_{22} \\ a_{32}a_{33} & a_{33}a_{31} & a_{31}a_{32} \end{matrix} \right\|$$

and define B_r similarly.

THEOREM 2. $\det B_r = -(\det A)^2(\det A_r)$.

Note that $\det A_r = (\Pi a_{ij})(\det \|1/a_{ij}\|)$ so that Theorem 1 is a consequence of Theorem 2 and the relation $(\operatorname{cof} A)/(\det A) = A^{-1}$.

LEMMA. $\det A_r = (\det A)a_{11}a_{22}a_{33} - b_{11}b_{22}b_{33}$.

Proof of Lemma. In the determinant of A_r, multiply the jth column by a_{1j}. Then subtract the first column from the others. This leads to

$$\begin{aligned} \det A_r &= \begin{vmatrix} a_{12}a_{23}a_{21} - a_{11}a_{22}a_{23} & a_{13}a_{21}a_{22} - a_{11}a_{22}a_{23} \\ a_{12}a_{33}a_{31} - a_{11}a_{32}a_{33} & a_{13}a_{31}a_{32} - a_{11}a_{32}a_{33} \end{vmatrix} \\ &= \begin{vmatrix} -a_{23}b_{33} & a_{22}b_{23} \\ a_{33}b_{32} & -a_{32}b_{22} \end{vmatrix} = a_{23}a_{32}b_{22}b_{33} - a_{22}a_{33}b_{23}b_{32}. \end{aligned}$$

Now substitute $a_{23}a_{32} = a_{22}a_{33} - b_{11}$ and $b_{23}b_{32} = b_{22}b_{33} - (\det A)a_{11}$ (from the relation above for the double cofactor) to obtain the Lemma.

The proof of Theorem 2 now follows by applying the Lemma both to A and to B and again using the double cofactor relation:

$$\det B_r = (\det B)b_{11}b_{22}b_{33} - (\det A)^3 a_{11}a_{22}a_{33}.$$

Since $\det B = (\det A)^2$ we have

$$\begin{aligned}\det B_r &= (\det A)^2[b_{11}b_{22}b_{33} - (\det A)a_{11}a_{22}a_{33}]\\ &= -(\det A)^2(\det A_r).\end{aligned}$$

Problem: Can anything be done for $n>3$?

ON POLYNOMIAL EQUATIONS WITH COEFFICIENTS EQUAL TO THEIR ROOTS

PAUL R. STEIN, Los Alamos Scientific Laboratory

1. Introduction. The identity $(z-x_1)(z-x_2)\cdots(z-x_n)=z^n+\sum_1^n y_i z^{n-i}$ defines a transformation $y=T_n(x)$ of complex n-space onto itself with fixed points $x=T_n(x)$ determined by

$$\text{(T)}\qquad x_i = (-1)^i\sigma_i(x_1, \cdots x_n); \qquad i = 1, \cdots n.$$

Here σ_i are the elementary symmetric functions of the quantities $x_1, \cdots, x_n$ [1].

It is clear that (a) zero is a fixed point of T_n; (b) a fixed point of T_n with exactly $m\geqq 1$ nonzero components must be of the form $(x_1, \cdots, x_m, 0, \cdots, 0)$, where $(x_1, \cdots, x_m)$ is a fixed point of T_m with *all* components nonzero, and (c) for every such fixed point of T_m, $(x_1, \cdots, x_m, 0, \cdots, 0)$ is a fixed point of T_n, $n\geqq m$.

The determination of fixed points thus reduces to the case with all $x_i\neq 0$. The latter type are shown to exist in the *real* subspace if and only if $n\leqq 4$; in other words, there are no polynomial equations of degree $n\geqq 5$ with real coefficients all different from 0 such that their coefficients are equal to their roots.

The nonexistence of these fixed points is demonstrated in Section 2; the special cases $n\leqq 4$ are discussed in Section 3.

2. Suppose that (T) holds for some $n\geqq 5$, with all x_i real and different from zero. The first two equations of (T) yield

$$\sum_1^n x_i^2 = x_1^2 - 2x_2.$$

Hence

$$0 < t \equiv \sum_2^n x_i^2 = -2x_2$$

$$0 < s \equiv \sum_3^n x_i^2 = -x_2(x_2+2) = 1 - (x_2+1)^2,$$

and so $-2<x_2<0$, $0<t<4$, $0<s\leqq 1$, and $|x_i|<1$ for $i\geqq 3$. Hence

$$(1)\qquad |x_3|\cdots|x_{n-1}|<\left(\frac{s}{n-3}\right)^{(n-3)/2}\leqq 1/(n-3)^{(n-3)/2}.$$

From the first equation of (T),

$$2|x_1|\leqq|x_2|+\cdots+|x_n|\leqq(n-1)\left(\frac{t}{n-1}\right)^{1/2}<2(n-1)^{1/2},$$

while from the last equation of the set we have $|x_1\cdots x_{n-1}|=1$. Combining these inequalities we get

$$1\leqq 2|x_1|\cdot|x_3|\cdots|x_{n-1}|<2[(n-1)/(n-3)^{n-3}]^{1/2}.$$

This is false for $n\geqq 6$.

Now suppose that $n=5$. From (1) we have $|x_3|\,|x_4|<s/2\leqq\frac{1}{2}$; therefore

$$(2)\qquad |x_1||x_2|>2/s\geqq 2.$$

From the first equation of (T),

$$|-x_2-2x_1|\leqq\sum_3^5|x_i|\leqq 3(s/3)^{1/2}\leqq\sqrt{3}.$$

This shows that $x_1>0$. Otherwise, we would have $|x_2|+2|x_1|\leqq\sqrt{3}$, implying

$$|x_1||x_2|\leqq\left(\frac{|x_1|+|x_2|}{2}\right)^2<3/4,$$

which contradicts (2). Thus $x_1>0$. But then, by Newton's relations for power sums in terms of the elementary symmetric functions,

$$\sum_1^5 x_i^3=-x_1^3+3x_1x_2-3x_3,$$

or $2x_1^3-|x_2|^3+3x_1|x_2|=-3x_3-\sum_3^5 x_i^3$. This shows that $x_1<|x_2|$. Otherwise, in view of (2), we would have the contradiction:

$$6<2x_1^3-|x_2|^3+3x_1|x_2|\leqq 3|x_3|+\sum_3^5|x_i|^3<3|x_3|+s<4.$$

Hence, by (2), $2/s<x_1|x_2|<|x_2|^2$, where $s=-x_2(x_2+2)$. Letting $x=|x_2|$, where, as shown above, $0<x<2$, we then have

$$(3)\qquad 2<f(x)\equiv x^3(2-x).$$

It is easily shown that $f(x)$ has the maximum $f(3/2)=27/16$ on the range in question. Thus (3) leads to a contradiction and the result is established.

3. The case $n=1$ is trivial. For $n=2$, we can solve the system (T) explicitly, obtaining the solution

$$x_1 = 1, \qquad x_2 = -2. \tag{4}$$

For $n=3$ the defining equation is $\lambda^3+x_1\lambda^2+x_2\lambda+x_3=0$. Substituting $x_1=\lambda$, and using $x_3=-2x_1-x_2$, we see that the equation factors;

$$(x_1 - 1)(2x_1^2 + 2x_1 + x_2) = 0. \tag{5}$$

It is easily verified that $x_1=1$ leads to the solution

$$x_1 = 1, \qquad x_2 = x_3 = -1. \tag{6}$$

If there is another solution with $x_1\neq 1$, then the second factor in (5) must vanish. Making use of the fact that $x_2=-(1/x_1)$, we find $2x_1^3+2x_1^2-1=0$, which leads to the second solution

$$x_1 = .56519772, \qquad x_2 = -1.76929234, \qquad x_3 = .63889690. \tag{7}$$

The case $n=4$ is treated similarly. Now $x_1=1$ again satisfies the defining equation identically; the latter reduces to

$$(x_1 - 1)[2x_1(1 + x_1 + x_1^2) - |x_2|(1 + x_1) + x_3] = 0. \tag{8}$$

Assuming that $x_1=1$, we derive for x_2 the equation

$$x_2^4 + 3x_2^3 + 3x_2^2 + 2x_2 + 1 = 0.$$

This has only one real root in the range $-2<x_2<0$, and we find the solution:

$$x_1 = 1, \quad x_2 = -1.7548782, \quad x_3 = -.5698401, \quad x_4 = .3247183. \tag{9}$$

To show that there are no further solutions, we must show that the second factor in (8) cannot vanish. In fact, the vanishing of this factor is equivalent to the equation $x_1(2x_1-|x_2|)+2x_1^3=x_4$. Clearly $2x_1>|x_2|$ leads to a contradiction, since it implies both $x_1>1/\sqrt{2}$ and $2x_1^3<x_4$; hence $x_4<|x_3|$, which yields $x_4^2+x_3^2>1$. Thus we would have $2x_1<|x_2|$. This, on the other hand, implies that $x_1<1/\sqrt{2}$, $|x_3|<x_4<2x_1^3<1/\sqrt{2}$. It would then follow that $x_1|x_3|<\frac{1}{2}$, requiring $|x_2|>2$, which is a contradiction. Thus there are no solutions with $x_1\neq 1$.

This paper was done under the auspices of the United States Atomic Energy Commission at the Los Alamos Scientific Laboratory of the University of California, Los Alamos, New Mexico

Reference

1. S. Ulam, A collection of Mathematical Problems, Interscience, New York, 1960, p. 30f.

A NOTE ON PROJECTIONS AND COMPACTNESS

S. B. NADLER, JR., University of Georgia

The purpose of this note is to give necessary and sufficient conditions for the projection functions on the cartesian product of certain types of spaces to be closed. Two theorems are proved showing, in special spaces, that a particular form of compactness is equivalent to closedness of the projection functions. The note is concluded with several examples illustrating that the apparently stringent hypotheses in Theorem 1 and Theorem 2 are necessary. It has been called to my attention that the first part of Theorem 1 is proved in [2].

If X_1 and X_2 are topological spaces, then $X_1 \times X_2 = \{(x_1, x_2): x_i \in X_i,\ i=1$ or $2\}$ is called the cartesian product of X_1 and X_2. The projection functions π_1 and π_2, from $X_1 \times X_2$ onto X_1 and X_2 respectively, are given by $\pi_i(x_1, x_2) = x_i$ for $i=1$ or 2.

A space X is said to be compact if and only if every cover by open sets has a finite subcover, sequentially compact if and only if every sequence has a convergent subsequence, countably compact if and only if every infinite subset has a limit point, and locally compact if and only if each point has a open neighborhood whose closure is compact. Finally, X satisfies the first axiom of countability if and only if, for each point p of X, there is a sequence $V_1^{(p)}, V_2^{(p)}, \cdots$ of open neighborhoods of p such that if U is any open set containing p, then there exists a set $V_i^{(p)}$ such that $V_i^{(p)} \subset U$. The reader is referred to [1] for definitions of terms not included here.

THEOREM 1. *Let X_1 and X_2 be nondiscrete T_1-spaces satisfying the first axiom of countability. Then X_1 is sequentially compact if and only if the projection function π_2, from $X_1 \times X_2$ onto X_2, is closed.*

Proof. The necessity follows from Lemma 1.1 in [2]. The condition is also sufficient. For suppose that π_2 is closed and let $\{x_i\}_{i=1}^{\infty}$ be a sequence of points of X_1. Let p be a limit point of X_2 and choose a sequence $\{y_i\}_{i=1}^{\infty}$ of distinct points of $X_2 - \{p\}$ such that $\lim_{i\to\infty} y_i = p$ (such a choice is possible since X_2 is a T_1-space satisfying the first axiom of countability). Let $S = \{(x_i, y_i): i=1, 2, \cdots\}$. If S had no limit point, then S would be closed. But this would imply that $\pi_2(S)$ is closed and, therefore, that $\{y_i: i=1, 2, \cdots\}$ is closed. This is a contradiction, showing that S has a limit point (x, y). Since X_1 and X_2 are T_1-spaces satisfying the first axiom of countability, so is $X_1 \times X_2$. Hence, there exists a subsequence $\{x_{i_j}\}_{j=1}^{\infty}$ of $\{x_i\}_{i=1}^{\infty}$ such that $\{(x_{i_j}, y_{i_j})\}_{j=1}^{\infty}$ is a sequence of points of S which converges to (x, y). Therefore, $\lim_{j\to\infty} x_{i_j} = x$ and the sequence $\{x_i\}_{i=1}^{\infty}$ has a convergent subsequence, showing that X_1 is sequentially compact.

Since countable compactness is equivalent to sequential compactness in T_1-spaces satisfying the first axiom to countability, countable compactness may be inserted in place of sequential compactness in Theorem 1. One notices that T_1 is necessary to the hypothesis for, if N denotes the natural numbers with a base for a topology consisting of $\{1, 2\}$ and all sets of the form $\{i\}$ for $i=3, 4, \cdots,$

then N is a first axiom space which is not countably compact thus not sequentially compact. But the projection functions on $N\times N\to N$ are closed.

THEOREM 2. *Let X_1 and X_2 be T_1-spaces, each having a countable base, such that neither X_1 nor X_2 is discrete. Then X_1 is compact if and only if the projection function π_2 is closed.*

Proof. The theorem is an immediate consequence of Theorem 1 since a T_1-space with a countable base is compact if and only if it is sequentially compact.

Since T_1, first axiom, countable base, and sequential compactness are each preserved by taking countable cartesian products, modifications of previous proofs yield generalizations of Theorem 1 and Theorem 2 to a countable number of spaces.

The requirement, in Theorem 1, that X_1 and X_2 both satisfy the first axiom of countability is necessary. To see this, let X_1 be the cartesian product of the unit interval $[0, 1]$ with itself an uncountable number of times and let X_2 be the unit interval. Since X_1 and X_2 are compact, π_i: $X_1\times X_2\to X_i$ (for $i=1$ or 2) is a mapping of a compact space onto a Hausdorff space and therefore is closed. However, X_1 is not sequentially compact, showing that closedness of the projection functions does not, in general, imply sequential compactness of the coordinate spaces.

The requirement of countable base in the hypothesis of Theorem 2 is necessary. Theorem 1 testifies to this, for there exist sequentially compact first axiom Hausdorff spaces which are not compact. For example, let X be the set of all ordinal numbers less than the first uncountable ordinal and let the topology on X be the order topology. It is well-known that X is a sequentially compact first axiom Hausdorff space which is locally compact but not compact. By Theorem 1, however, the projection functions on $X\times X$ are closed. Therefore, even for first axiom locally compact spaces, closedness of the projection functions does not imply compactness of the coordinate spaces.

References

1. J. L. Kelley, General Topology, Van Nostrand. Princeton, 1955.
2. Takesi Isiwata, Normality and perfect mappings, Proc. Japan Acad., 39 (1963) 95–97.

IDEMPOTENT MATRICES (mod p^a)

J. H. HODGES, University of Colorado

1. Introduction. In a mathematical note several years ago, Irma Reiner [**3**, Section 4] gave a formula for the number of, as well as an explicit construction for, all matrix solutions of order m of $X^2\equiv I$ (mod p^a), where p is an odd prime and a is an arbitrary positive integer. The case $a=1$ is a special case of an earlier result [**2**] for matrices over an arbitrary *finite* field. Recently, Donna F. Rochon [**4**; Ch. 2] has shown that the recursive method used by Mrs. Reiner

in her note leads to some special difficulties, as yet unresolved, when applied to the matrix congruence $X^2 \equiv I \pmod{2^a}$ with $a > 1$.

In the present note, it is shown that this same recursive method can be applied to give the number of matrix solutions of order m of $X^2 \equiv X \pmod{p^a}$ for arbitrary prime p and positive integer a. Again, for $a=1$, the result obtained here is a special case of a previous result [1] concerning matrix solutions of scalar polynomial equations over a finite field.

2. Idempotent matrices $\pmod{p^a}$. Let a and m be arbitrary positive integers and p be an arbitrary prime. It can be shown easily [1; p. 291] that a matrix $Q \pmod{p}$ of order m satisfies $Q^2 \equiv Q \pmod{p}$ if and only if Q is similar $\pmod{p}$ to some $C_t = \operatorname{diag}(I_t, 0)$ for $0 \leqq t \leqq m$, where I_t is the identity matrix of order t.

Now, any solution X of order m of

$$X^2 \equiv X \pmod{p^{a+1}} \tag{1}$$

is congruent mod p^a to some solution R of order m of

$$X^2 \equiv X \pmod{p^a}, \tag{2}$$

where R in turn is congruent $\pmod{p}$ to some solution Q of order m of

$$X^2 \equiv X \pmod{p}. \tag{3}$$

Let R be a solution of order m of (2) such that $R \equiv C_t \pmod{p}$ for some fixed t, $0 \leqq t \leqq m$. In order to find all solutions of (1) which are congruent to $R \pmod{p^a}$, we let $X = R + p^a T$ and find all matrices $T \pmod{p}$ such that $(R+p^aT)^2 \equiv (R+p^aT) \pmod{p^{a+1}}$. Equivalently, we find all T such that

$$C_tT + TC_t - T \equiv (R - R^2)/p^a \pmod{p}. \tag{4}$$

If T of order m is partitioned as $T = [T_{ij}]$ for $1 \leqq i, j \leqq 2$, where T_{11} is $t \times t$, T_{12} is $t \times (m-t)$, T_{21} is $(m-t) \times t$ and T_{22} is $(m-t) \times (m-t)$, then (4) may be written as

$$\begin{bmatrix} T_{11} & -T_{12} \\ -T_{21} & T_{22} \end{bmatrix} \equiv (R - R^2)/p^a \pmod{p}. \tag{5}$$

Now (5) and so (4) is uniquely solvable for all the submatrices T_{ij} so that there is a unique solution of (1) which is congruent to this solution R of (2). Thus, by induction, there exists a unique solution X of order m of (1) such that $X \equiv C_t \pmod{p}$.

Let t be a fixed integer, $0 < t < m$ and P be an integral matrix with determinant prime to p. Then P has an inverse $P^{-1} \pmod{p^a}$ for each a. If now X is the solution of (1) which is congruent to $C_t \pmod{p}$, then $P^{-1}XP \pmod{p^{a+1}}$ is a solution of (1) which is congruent to $Q = P^{-1}C_tP \pmod{p}$.

Furthermore, if Q_1 and Q_2 are incongruent solutions $\pmod{p}$ of (3) then their associated solutions X_1 and X_2 of (1) are unique and incongruent $\pmod{p^{a+1}}$. Therefore, there are exactly as many solutions X of order m of (1)

which are incongruent (mod p^{a+1}) (and also solutions R of order m of (2) which are incongruent (mod p^a)) as there are solutions Q of order m of (3) which are incongruent (mod p). This latter number, which we obtain by taking $q=p$ and $E(x)=x^2-x=x(x-1)$ in Theorem 2 of [1], is equal to

$$g_m \sum_{t=0}^{m} (g_t g_{m-t})^{-1}, \tag{6}$$

where $g_0=1$ and, for $t>0$, g_t is the number of nonsingular matrices of order t (mod p), which is given by

$$g_t = \prod_{i=0}^{t-1} (p^t - p^i). \tag{7}$$

The work on this paper has been supported by National Science Foundation Research Grant 19894.

References

1. J. H. Hodges, Scalar polynomial equations for matrices over a finite field, Duke Math. J., 25 (1958) 291–296.

2. ———, The matrix equation $X^2-I=0$ over a finite field, this MONTHLY, 65 (1958) 518–520.

3. Irma Reiner, The matrix congruence $X^2\equiv I$ (mod p^a), this MONTHLY, 67 (1960) 773–775.

4. Donna F. Rochon, Some matrix congruences over the ring of integers modulo p^a, unpublished M.A. thesis, University of Colorado, June, 1963.

ON EXISTENCE THEOREMS FOR DIFFERENCE AND q-DIFFERENCE EQUATIONS

SELMO TAUBER, Portland State College

1. Introduction. The operators $E(h)$ and $A(q)$ are defined, respectively, by the relations

$$E(h)f(t) = f(t+h), \quad E^2(h)f(t) = f(t+2h), \cdots, E^n(h)f(t) = f(t+nh),$$
$$A(q)f(t) = f(qt), \quad A^2(q)f(t) = f(q^2t), \cdots, A^n(q)f(t) = f(q^nt).$$

In particular, for $h=1$, $E(1)=E$, thus, $E^nf(t)=f(t+n)$.

In [1], for example, the uniqueness of the solutions of the difference equation

$$E^nx(t) = \phi[t, x(t), Ex(t), E^2x(t), \cdots, E^{n-1}x(t)]$$

is proved assuming that a solution exists for $t_0\leq t\leq t_0+1$. In [2] this is proved more generally for a system of simultaneous difference equations in normal form

$$E^{n_j}x_j(t) = \phi_j[t, x_1(t), Ex_1(t), \cdots, E^{n_1-1}x_1(t), x_2(t), Ex_2(t), \cdots, E^{n_2-1}x_2(t), \cdots,$$
$$x_m(t), Ex_m(t), \cdots, E^{n_m-1}x_m(t)], \qquad (j = 1, 2, \cdots, m)$$

by writing it in the form of a vector difference equation

$$E\mathbf{Y} = \mathbf{F}(\mathbf{Y}, t), \tag{1}$$

where it is assumed that for $t_0 \leqq t \leqq t_0+1$, $\mathbf{Y}=\mathbf{Y}_0$, and $\mathbf{Y}_0$ is known for $t_0 \leqq t \leqq t_0+1$.

In [3] the same is proved for a system of simultaneous q-difference equations, assuming that $q>1$ the system of simultaneous q-difference equations can be written

$$A^{n_j}(q)x_j(t) = \phi_j[t, x_1(t), A(q)x_1(t), \cdots, A^{n_1-1}(q)x_1(t), x_2(t), A(q)x_2(t), \cdots, A^{n_2-1}(q)x_2(t), \cdots, x_m(t), A(q)x_m(t), \cdots, A^{n_m-1}(q)x_m(t)],$$
$$j = 1, 2, \cdots, m.$$

This system of simultaneous q-difference equations can be written in the form of a vector q-difference equation

$$A(q)\mathbf{Y} = \mathbf{F}(\mathbf{Y}, t), \tag{2}$$

under the same assumptions as for the vector difference equation (1). More about q-difference equations, especially about linear q-difference equations, can be found in [4], where extensive reference to the literature is given.

In this note we shall prove that

(i) equation (1) has a unique solution if for $t=t_0$, $\mathbf{Y}=\mathbf{Y}_0(t_0)$, and if (1) satisfies condition $(\mathfrak{C}_1)$, defined below;

(ii) equation (2) has a unique solution if for $t=t_0$, $\mathbf{Y}=\mathbf{Y}_0(t_0)$, and if (2) satisfies condition $(\mathfrak{C}_2)$, defined below.

2. Case of difference equations. We consider the difference equation

$$E(h)\mathbf{Y} = \mathbf{f}(\mathbf{Y}, t, h), \tag{3}$$

which is defined for $0<h\leqq 1$. If (3) is defined for $h=0$, then

$$\mathbf{Y} = \mathbf{f}(\mathbf{Y}, t, 0) = \mathbf{G}(\mathbf{Y}, t).$$

For $h=1$, (3) becomes $E\mathbf{Y}=\mathbf{f}(\mathbf{Y}, t, 1)=\mathbf{F}(\mathbf{Y}, t)$, i.e. (3) becomes (1). Conversely, given (1), it is possible to write an equation of form (3) reducing to (1) for $h=1$. One such equation would be, $E(h)\mathbf{Y}=\mathbf{F}(\mathbf{Y}, s/h)$, obtained from (1) by changing t into s/h, but this is not the only one. We shall call (3) the *h-associate* of (1). We define condition $(\mathfrak{C}_1)$ as follows:

Equation (1) is said to satisfy condition $(\mathfrak{C}_1)$ if there exists an h-associate of (1) defined for all h such that $0<h\leqq 1$.

We assume in addition that for $t=t_0$, $\mathbf{Y}=\mathbf{Y}_0(t_0)=[C_1, C_2, \cdots, C_n]$, where $C_k=y_k(t_0)$, $k=1, 2, \cdots, n$. Thus

$$E(h)\mathbf{Y} = [y_1(t_0+h), y_2(t_0+h), \cdots, y_n(t_0+h)] = \mathbf{Y}(h) = \mathbf{f}(\mathbf{Y}_0, t_0, h) = \mathbf{f}(h).$$

This defines $\mathbf{Y}$ for $t_0\leqq t\leqq t_0+1$, thus $\mathbf{Y}_0(t)$. We are therefore in the conditions of the existence and uniqueness theorems of [1] and [2] and can state:

THEOREM I. *If for* $t=t_0$, $\mathbf{Y}=\mathbf{Y}_0(t_0)=[C_1, C_2, \cdots, C_n]$, *and if it satisfies condition* $(\mathcal{C}_1)$, *then the vector difference equation* $E\mathbf{Y}=\mathbf{F}(\mathbf{Y}, t)$ *has a unique solution for* $t\geqq t_0$.

3. Case of q-difference equations. We consider (2) which we can write

$$A(q)\mathbf{Y} = \mathbf{Y}(qt) = \mathbf{F}[\mathbf{Y}(t), t] = \mathbf{f}[\mathbf{Y}(t), t, q].$$

Let us consider q as a given number $q>1$, and let $1<p\leqq q$. The equation

$$A(p)\mathbf{Y} = \mathbf{Y}(pt) = \mathbf{f}[\mathbf{Y}(t), t, p] \tag{4}$$

is said to be the p-associate of (2). We define condition $(\mathcal{C}_2)$ as follows:

Equation (2) is said to satisfy condition $(\mathcal{C}_2)$ if there exists a p-associate of (2) defined for any p such that $1<p\leqq q$. In addition we assume the boundary conditions

$$\mathbf{Y}(t_0) = \mathbf{Y}_0(t_0) = [y_1(t_0), y_2(t_0), \cdots, y_n(t_0)] = [C_1, C_2, \cdots, C_n].$$

It follows that $\mathbf{Y}(pt_0)=\mathbf{f}[\mathbf{Y}(t_0), t_0, p]$, or $\mathbf{Y}(p)=\mathbf{f}(p)$, which defines $\mathbf{Y}_0=\mathbf{Y}=[y_1(t), y_2(t), \cdots, y_n(t)]$ for $t_0\leqq t\leqq qt_0$. We are therefore in the situation of [3] and can apply the results obtained there. We state:

THEOREM II. *If for* $t=t_0$, $\mathbf{Y}=\mathbf{Y}_0(t_0)=[C_1, C_2, \cdots, C_n]$, *and if it satisfies condition* $(\mathcal{C}_2)$, *then the vector q-difference equation* $A(q)\mathbf{Y}=\mathbf{F}[\mathbf{Y}(t), t]$ *has a unique solution for* $t\geqq t_0$.

References

1. K. S. Miller, An Introduction to the Calculus of Finite Differences and Difference Equations, Holt, Rinehart and Winston, New York, 1960.

2. S. Tauber, Existence and uniqueness theorems for solutions of difference equations, this MONTHLY, 71 (1964) 859–862.

3. ———, Sur les Opérateurs A et B, Elem. Math., 20 (1965) 81–87.

4. C. R. Adams, Linear q-difference equations, Bull. Amer. Math. Soc., 37 (1931) 361–400.

ON THE TARRY-ESCOTT PROBLEM

T. N. SINHA, Bhagalpur University, Bihar (India)

The Tarry-Escott problem is that of finding two sets of integers, which may be assumed (since zero is allowed) to be equal in number, such that the integers in each set have the same sum, the same sum of squares, etc., up to and including the same sum of the kth powers, i.e., we are to find solutions in integers of the system of equations

$$\sum_{i=1}^{s} a_i^x = \sum_{i=1}^{s} b_i^x \qquad (x = 1, 2, 3, \cdots, k). \tag{1}$$

This system will be denoted here by the notation

$$a_1, a_2, \cdots, a_s \overset{x}{=} b_1, b_2, \cdots, b_s \qquad (x = 1, 2, \cdots, k).$$

The problem has attracted the attention of number theorists since the time of Goldbach and Euler, and has several interesting applications. A number of authors, expecially recently, have been interested in finding the least value of s for which there exists a solution of (1). The following easy theorem was first established by L. Bastein [**1**].

THEOREM 1. *If the equations* (1) *have a nontrivial solution,* (*i.e.,* $\{b_i\}$ *is not a permutation of* $\{a_i\}$) *then* $s \geqq k+1$.

Solutions of (1) with $s=k+1$ have been called ideal solutions by J. Chernick [**2**]. Ideal solutions of (1) have so far been obtained only for $k \leqq 9$. For $k=8$ and $k=9$, ideal solutions of (1) in parametric forms are still wanted.

In this note I indicated a parametric solution of the system of equations

$$\text{(2)} \qquad A_1, A_2, A_3, A_4, A_5 \overset{x}{=} B_1, B_2, B_3, B_4, B_5 \qquad (x = 1, 3, 5, 7).$$

This implies the system (1) for $k=8$ and $s=10$ in view of the following well-known

THEOREM 2. *If* $a_1, a_2, \cdots, a_p \overset{x}{=} b_1, b_2, \cdots, b_p (x=1, 3, 5, \cdots, 2n-1)$ *then*

$$t + a_1, t + a_2, \cdots, t + a_p, t - b_1, t - b_2, \cdots, t - b_p$$
$$\overset{x}{=} t + b_1, t + b_2, \cdots, t + b_p, t - a_1, t - a_2, \cdots, t - a_p \quad (x = 1, 2, 3, \cdots, 2n)$$

where t is arbitrary.

To solve (2) we shall use the following

THEOREM 3. *If* $a_1, a_2, a_3 \overset{x}{=} b_1, b_2, b_3$ $(x=2, 4)$ *and*

$$a_1 + a_2 - a_3 = 2(b_1 + b_2 - b_3), \qquad a_3 \neq a_1 + a_2, \qquad b_3 \neq b_1 + b_2$$

then

$$2a_1 - 3h, 2a_1 - h, 2a_2 - 3h, 2a_2 - h, 2b_3 + h$$
$$\overset{x}{=} 2a_3 + h, 2a_3 + 3h, 2b_1 - h, 2b_2 - h, 3h \ (x = 1, 3, 5, 7),$$

where $2h = b_1 + b_2 - b_3$.

Proof of Theorem 3. Set $s_2 = a_1^2 + a_2^2 + a_3^2$, $t_2 = b_1^2 + b_2^2 + b_3^2$, $s_4 = a_1^4 + a_2^4 + a_3^4$, $t_4 = b_1^4 + b_2^4 + b_3^4$ and let

$$f_n(s_2, s_4) = (a_1 + a_2 + a_3)^n + (-a_1 + a_2 + a_3)^n + (a_1 - a_2 + a_3)^n$$
$$+ (a_1 + a_2 - a_3)^n - (2a_1)^n - (2a_2)^n - (2a_3)^n.$$

We find

$$f_1(s_2, s_4) = 0, \qquad f_2(s_2, s_4) = 0, \qquad f_4(s_2, s_4) = 12(s_2^2 - 2s_4)$$

$$f_6(s_2, s_4) = 15s_2(s_2^2 - 2s_4), \qquad f_8(s_2, s_4) = 28(7s_2^2 + 2s_4)(s_2^2 - 2s_4).$$

Since, by hypothesis, $s_2 = t_2$, $s_4 = t_4$ we have $f_n(s_2, s_4) = f_n(t_2, t_4)$ for $n = 1, 2, 4, 6, 8$, i.e.

$$\text{(A)} \quad \begin{aligned} &2a_1, 2a_2, b_1 + b_2 + b_3, 2a_3, b_1 - b_2 + b_3, -b_1 + b_2 + b_3, b_1 + b_2 - b_3 \\ &\overset{x}{=} a_1 - a_2 + a_3, -a_1 + a_2 + a_3, 2b_3, a_1 + a_2 + a_3, 2b_1, 2b_2, a_1 + a_2 - a_3 \\ &\qquad\qquad (x = 1, 2, 4, 6, 8). \end{aligned}$$

Now since

$$\text{(A}_0\text{)} \qquad a_1 + a_2 - a_3 = 2(b_1 + b_2 - b_3) = 4h$$

by hypothesis, we can write (A) as

$$\text{(B)} \quad \begin{aligned} &(2a_1-h)+h, (2a_2-h)+h, (2b_3+h)+h, (2a_3+h)-h, (2b_1-h)-h, (2b_2-h)-h, 3h-h \\ &\overset{x}{=} (2a_1-3h)-h, (2a_2-3h)-h, (2b_3+h)-h, (2a_3+3h)+h, (2b_1-h)+h, \\ &\qquad (2b_2-h)+h, 3h+h \qquad (x = 1, 2, 4, 6, 8). \end{aligned}$$

Let $A_1 = 2a_1 - 3h$, $A_2 = 2a_1 - h$, $A_3 = 2a_2 - 3h$, $A_4 = 2a_2 - h$, $A_5 = 2b_3 + h$, $B_1 = 2a_3 + h$, $B_2 = 2a_3 + 3h$, $B_3 = 2b_1 - h$, $B_4 = 2b_2 - h$, $B_5 = 3h$. Then $A_1 + h = A_2 - h$, $A_3 + h = A_4 - h$, $B_1 + h = B_2 - h$, while $A_1 + A_2 + A_3 + A_4 + A_5 = B_1 + B_2 + B_3 + B_4 + B_5$ by (A_0). Hence (B) implies

$$\text{(C)} \quad \begin{aligned} &h+A_1, h+A_2, h+A_3, h+A_4, h+A_5, h-B_1, h-B_2, h-B_3, h-B_4, h-B_5 \\ &\overset{x}{=} h+B_1, h+B_2, h+B_3, h+B_4, h+B_5, h-A_1, h-A_2, h-A_3, h-A_4, h-A_5 \\ &\qquad\qquad (x = 1, 2, 4, 6, 8). \end{aligned}$$

Expanding the terms of the above identity by the binomial theorem for each power x in succession, it is readily seen that (C) implies

$$A_1, A_2, A_3, A_4, A_5 \overset{x}{=} B_1, B_2, B_3, B_4, B_5 \qquad (x = 1, 3, 5, 7)$$

which proves the theorem. We therefore seek to solve the system

$$\text{(3)} \quad \begin{aligned} a_1^2 + a_2^2 + a_3^2 &= b_1^2 + b_2^2 + b_3^2 \\ a_1^4 + a_2^4 + a_3^4 &= b_1^4 + b_2^4 + b_3^4 \\ a_1 + a_2 - a_3 &= 2(b_1 + b_2 - b_3). \end{aligned}$$

We shall first prove the

LEMMA. *If the equations* (3_1) *and* (3_2) *hold, then*

$$\text{(4)} \qquad (b_3^2 - a_1^2)(b_3^2 - a_2^2) = (a_3^2 - b_1^2)(a_3^2 - b_2^2).$$

Proof of the Lemma. Squaring (3_1) and subtracting (3_2) we get

$$a_1^2a_2^2 + a_2^2a_3^2 + a_3^2a_1^2 = b_1^2b_2^2 + b_2^2b_3^2 + b_3^2b_1^2. \tag{4'}$$

Now

$$\begin{aligned}(b_3^2 - a_1^2)(b_3^2 - a_2^2) &= (a_2^2 + a_3^2 - b_1^2 - b_2^2)(a_1^2 + a_3^2 - b_1^2 - b_2^2)\\ &= (a_2^2 + a_3^2)(a_1^2 + a_3^2) - (b_1^2 + b_2^2)(a_1^2 + a_2^2 + 2a_3^2 - b_1^2 - b_2^2)\\ &= (a_2^2 + a_3^2)(a_1^2 + a_3^2) - (b_1^2 + b_2^2)(b_3^2 + a_3^2), &&\text{by } (3_1)\\ &= a_3^4 - a_3^2(b_1^2 + b_2^2) + a_1^2a_2^2 + a_2^2a_3^2 + a_3^2a_1^2 - b_1^2b_3^2 - b_2^2b_3^2\\ &= a_3^4 - a_3^2(b_1^2 + b_2^2) + b_1^2b_2^2, &&\text{by}(4')\\ &= (a_3^2 - b_1^2)(a_3^2 - b_2^2),\end{aligned}$$

which proves the lemma.

Conversely, one would easily see that the equations (3_1) and (4) imply the equation (3_2). Hence, we need consider only the equations (3_1), (3_3) and (4). To satisfy (3_3), set

$$a_3 = a_1 + b_1 - b_2, \qquad 2b_3 = 3b_1 + b_2 - a_2. \tag{5}$$

Then, substituting in (3_1), we have

$$4a_1^2 + 4a_2^2 + 4(a_1 + b_1 - b_2)^2 = 4b_1^2 + 4b_2^2 + (3b_1 + b_2 - a_2)^2,$$

which can be written as

$$\begin{aligned}4a_1^2 + 4a_2^2 + 4(a_1 + b_1)^2 - 8b_2(a_1 + b_1) + 4b_2^2&\\ = 4b_1^2 + 4b_2^2 + (3b_1 + b_2 + a_2)^2 - 4a_2(3b_1 + b_2 + a_2) + 4a_2^2&\end{aligned}$$

or

$$4(a_1 + b_1)(a_1 - b_1) + 4(a_1 + b_1)^2 - 8b_2(a_1 + b_1) = (3b_1 + b_2 + a_2)(3b_1 + b_2 - 3a_2),$$

i.e.,

$$8(a_1 + b_1)(a_1 - b_2) = (3b_1 + b_2 + a_2)(3b_1 + b_2 - 3a_2). \tag{6}$$

Again from (4), substituting for a_3 and b_3, we get

$$\begin{aligned}&16(a_1 - b_2)(a_1 + 2b_1 - b_2)(a_1 + b_1 - 2b_2)(a_1 + b_1)\\ &\quad = (3b_1 + b_2 - a_2 - 2a_1)(3b_1 + b_2 - a_2 + 2a_1)(3b_1 + b_2 - 3a_2)(3b_1 + b_2 + a_2)\end{aligned}$$

or by (6)

$$2(a_1 + 2b_1 - b_2)(a_1 + b_1 - 2b_2) = (3b_1 + b_2 - a_2 - 2a_1)(3b_1 + b_2 - a_2 + 2a_1). \tag{7}$$

Thus, if the equations (3_1), (3_3) and (4) hold, then equations (5), (6) and (7) hold. Conversely, it can be seen readily that the system of equations (5), (6)

and (7) imply the system (3_1), (3_3) and (4), and hence the system (3). To obtain a parametric solution of the system (5), (6) and (7) we proceed as follows:

Eliminating a_1 between (6) and (7) we get

$$16(b_1^2 - b_1b_2 + b_2^2) = (3b_1 + b_2 - a_2)^2 + 12a_2^2$$

or, $(13a_2+7b_1-15b_2)(a_2-b_1+b_2)=0$. If $a_2-b_1+b_2=0$, then $a_3=a_1+a_2$, $b_3=b_1+b_2$, which case is excluded in Theorem 3. Hence,

$$(8) \qquad 13a_2 + 7b_1 - 15b_2 = 0.$$

Again, eliminating a_2 between (7) and (8) we get

$$169a_1^2 + 169a_1(b_1 - b_2) - 240b_1^2 - 251b_1b_2 + 112b_2^2 = 0,$$

whence

$$(9) \qquad a_1 = [-169(b_1 - b_2) \pm 13\alpha]/2\cdot 169,$$

where

$$(10) \qquad \alpha^2 = 1129b_1^2 + 666b_1b_2 - 279b_2^2.$$

A simple parametric solution of (10) can be found as follows: A simple numerical solution of (10) is $b_1=-7$, $b_2=8$, $\alpha=13$. The equation (10) can be written as

$$1129\alpha^2 = (1129b_1 + 333b_2)^2 - Db_2^2,$$

where $D=333^2+1129\cdot 279$ so that, $1129\cdot 13^2=[1129(-7)+333\cdot 8]^2-\mathrm{D}\cdot 8^2$.

Eliminating D between the last two equations, we get

$$(8\alpha + 13b_2)(8\alpha - 13b_2) = (9032b_1 - 2575b_2)(8b_1 + 7b_2).$$

Set $p/q=(8\alpha-13b_2)/(8b_1+7b_2)$, where p and q are integers. We thus get the equations

$$8pb_1 + (7p + 13q)b_2 - 8q\alpha = 0,$$

and $9032qb_1-(2575q+13p)b_2-8p\alpha=0$, which give (discarding proportions)

$$b_1 = -7p^2 - 26pq - 2575q^2,\ b_2 = 8(p^2 - 1129q^2), \alpha = -13(p^2 + 806pq + 1129q^2).$$

To obtain simpler forms for b_1, b_2, α, we use the transformation

$$p = 16m - 13n, \qquad q = -n,$$

where m and n are integers. We thus obtain, discarding the common factor 256,

$$b_1 = -7m^2 + 13mn - 16n^2, \qquad b_2 = 8m^2 - 13mn - 30n^2,$$

$$\alpha = 13(-m^2 + 52mn - 46n^2)$$

as a parametric solution of (10). Hence from (9), taking the plus sign before α,

$$a_1 = 7m^2 + 13mn - 30n^2.$$

Then from (8), $a_2 = 13m^2 - 22mn - 26n^2$. Finally from (5),

$$a_3 = -8m^2 + 39mn - 16n^2, \qquad b_3 = -13m^2 + 24mn - 26n^2.$$

The negative sign before α only interchanges a_1 and a_3 with sign changed. If we denote the quadratic form $am^2+bmn+cn^2$ by the notation $[a, b, c]$, we write the solution of the system (3) as

$$a_1 = [7, 13, -30], \qquad a_2 = [13, -22, -26], \qquad a_3 = [-8, 39, -16]$$
$$b_1 = [-7, 13, -16], \qquad b_2 = [8, -13, -30], \qquad b_3 = [-13, 24, -26].$$

By Theorem 3, the system (2) has then the following parametric solution:

$$A_1 = [-7, 62, -30], \qquad A_2 = [7, 38, -50], \qquad A_3 = [5, -8, -22],$$
$$A_4 = [19, -32, -42], \qquad A_5 = [-19, 36, -62], \qquad B_1 = [-9, 66, -42],$$
$$B_2 = [5, 42, -62], \qquad B_3 = [-21, 38, -22], \qquad B_4 = [9, -14, -50],$$
$$B_5 = [21, -36, -30].$$

References

1. L. Bastein, Sphinx-Oedipe, 8 (1913) 171–172.
2. J. Chernick, Ideal solutions of the Tarry-Escott problem, this MONTHLY, 44 (1937) 626–633.
3. Gazeta Matematica, 48 (1942) 68–69.

PROJECTING m ONTO c_0

ROBERT WHITLEY, New Mexico State University

It is a well-known result, due to Phillips, that the Banach space m, of bounded sequences with the sup norm, cannot be projected continuously onto the subspace c_0 of sequences converging to zero [1, page 33, Corollary 4]. A typical use of this fact is found in [2]. We give a simple proof using an idea inherent in [4] and, as was pointed out by the referee, in [3]. Our method may also be used to simplify the proof of the result in [4].

LEMMA [5, page 77]. *Let I be a countable set. Then there is a family $\{U_a: a \text{ in } A\}$ of subsets of I such that* (1) *U_a is infinite,* (2) *$U_a \cap U_b$ is finite for $a \neq b$ and* (3) *the index set A is uncountable.*

Proof. Arthur Kruse has given the following elegant proof: Take I to be the rationals in (0, 1), A the irrationals in (0, 1) and, for a in A, let U_a be a sequence of rationals in (0, 1) converging to a.

Recall that a subset of the conjugate space X^* of a Banach space X is total if the only vector annihilated by all members of the subset is the zero vector.

For brevity we say that a Banach space X has (property) B if X^* contains a countable total subset. It is easy to see that B is preserved under isomorphism, that a subspace of a space with B has B and that the space m has B.

THEOREM. *There is no continuous projection of m onto c_0.*

Proof. Suppose that there is a continuous projection of m onto c_0. Then $m=c_0\oplus R$, where R is a closed subspace of m. Since m/c_0 is isomorphic to R we see that m/c_0 has B. The proof consists of showing that m/c_0 does not have B.

We think of m as $B(I)$, the bounded functions on a countable set I. Let $\{U_a: a \text{ in } A\}$ be a family of subsets of I as in the lemma and let f_a be the coset in m/c_0 which contains the characteristic function of the set U_a.

Let g be in $(m/c_0)^*$. We will show that the set $\{f_a: g(f_a)\neq 0\}$ is countable; it suffices to show that the set $C(n)=\{f_a: |g(f_a)| \geqq 1/n\}$ is countable for each natural number n. Choose $f_1, \cdots, f_m$ in $C(n)$ and let $b_i=\operatorname{sgn}(g(f_i))=\overline{g(f_i)}/|g(f_i)|$. The vector $x=\sum b_i f_i$ is of norm one (note that as a coset x contains vectors whose norm may be greater than one), and so $\|g\| \geqq |g(x)| \geqq m/n$; thus $C(n)$ is finite for each n.

We conclude by noting that if $\{h_i\}$ is a countable subset of $(m/c_0)^*$ then our argument shows that there are only countably many f_a with $h_i(f_a)$ nonzero for some i. Hence we can find a vector f_a which is mapped into zero by all the h_i, and so the set $\{h_i\}$ is not total.

References

1. M. M. Day, Normed Linear Spaces, Academic Press, New York, 1962.
2. J. Lindenstrauss, A remark concerning projections in summability domains, this MONTHLY, 70 (1963) 977–978.
3. M. Nakamura and S. Kakutani, Banach limits and the Čech compactification of a countable discrete set, Proc. Imp. Acad. Japan, 19 (1943) 224–229.
4. A. Pełczyński and V. N. Sudakov, Remark on non-complemented subspaces of the space $m(S)$, Colloq. Math., 9 (1962) 85–88.
5. W. Sierpinski, Cardinal and Ordinal Numbers, Warszawa, 1958.

INTERIORITY AND THE TONELLI CONDITIONS

W. V. CALDWELL, Flint College, Univ. of Michigan

In 1937, S. Stoïlow proved that if f is a complex-valued function of a complex variable which has the properties: (i) point inverses are totally disconnected, and (ii) f maps interior points of its domain of definition into interior points of the image, then f is topologically equivalent to an analytic function. This result stimulated interest in light interior functions (i.e. functions satisfying (i) and (ii)) and in establishing conditions which insure that a function satisfying these conditions will be light and interior. Titus and Young proved that if $f\in C'$ and

the Jacobian determinant of f is nonnegative and vanishes on at most a totally disconnected set, then f is light, interior, and orientation-preserving.

These conditions are rather strong and in weakening them one naturally considers the Tonelli conditions. Evans has shown that if u is absolutely continuous in the sense of Tonelli (ACT), then u is a continuous "potential function of its generalized derivatives" and conversely. If, however, u is only of bounded variation in the sense of Tonelli (BVT), then u will have actual partial derivatives a.e. but these derivatives will not in general equal the generalized (distribution) derivatives of u.

Consider an elliptic system $\mathfrak{L}$ given by the equations

$$\begin{aligned} U_x &= aV_x + bV_y \\ -U_y &= cV_x + dV_y \end{aligned} \qquad 4bc - (a+d)^2 \geqq \epsilon > 0,\ b > 0, \tag{1}$$

where a, b, c, and d are uniformly bounded real-valued measurable functions defined in a domain $\mathfrak{D}$ in the complex plane. Morrey and Bers and Nirenberg have investigated the properties and the representation of functions $f=u+iv$ which are ACT and which satisfy (1) wherever u and v have partial derivatives. In particular, it was shown that under these assumptions, f is light, interior, and orientation-preserving. The purpose of this note is to give an example which shows that the condition that f be ACT is essential.

Let $\mathfrak{D}=\{(x, y) \mid 0\leqq x\leqq 1, 0\leqq y\leqq 1\}$, let p be a continuous real-valued strictly increasing function defined on $[0, 1]$ such that (i) p has a derivative a.e., (ii) $p'=0$ wherever the derivative exists, and let q be defined similarly. Then $g(x, y)=p(x)+iq(y)$ is a homeomorphism. Now let ψ be the harmonic function defined in $\mathfrak{D}$ and such that $\psi+q$ vanishes on the boundary of $\mathfrak{D}$ and let ϕ be the harmonic conjugate of ψ. Then if $u=\phi+p$ and $v=\psi+q$, u and v have partial derivatives a.e. in $\mathfrak{D}$ and $f=u+iv$ satisfies the Cauchy-Riemann equations wherever these partials exist. Clearly, f is not interior.

References

1. L. Bers and L. Nirenberg, On a representation theorem for linear elliptic systems with discontinuous coefficients and its applications, *Convegno Internazionale Sulle Equazioni Lineari Alle Derivate Parziali*, Trieste, 25–38 Agosto, 1954.

2. G. C. Evans, Complements of potential theory, II. Am. J. Math., 55 (1933) 29–49.

3. G. C. Evans, Fundamental points of potential theory, Rice Institute Pamphlets, **7**, no. 4 (1920) 252–329.

4. K. Krickeberg, Distribution, Funktionen beschränkter Variation und Lebesguescher Inhalt nichtparametrischer Flächen, Ann. Mat. Pura App., 44 (1957) 105–133.

5. C. B. Morrey, On the solutions of quasi-linear elliptic partial differential equations, Trans. Amer. Math. Soc., **43** (1938) 126–166.

6. C. J. Titus and G. S. Young, A Jacobian condition for interiority, Mich. Math. J. 1 (1952) 89–93.

A NON-ENUMERABLE EXCEPTIONAL SET

D. Borwein, University of Western Ontario

Let $\{x_n\}$ be a sequence of real positive numbers not converging to 0 and let A be the set of all real numbers a for which $\{x_n\}$ converges to 0 (mod a). The problem posed in this Journal of showing that the exceptional set A must be of measure 0 was solved by I. J. Schoenberg [1964, 332] who asked whether or not A is necessarily enumerable. In this note I show that A need not be enumerable.

Let $x_n = 2^{n-1}(n-1)!$ and, for $n=1, 2, \cdots; k=0, \pm 1, \cdots$, let $I(k, n)$ be the closed interval

$$\left[\frac{k-\frac{1}{4n}}{x_n}, \frac{k+\frac{1}{4n}}{x_n}\right].$$

Then $I(k, n)$ contains the three intervals $I(j, n+1), j=2nk-1, 2nk, 2nk+1$, all other intervals $I(j, n+1)$ being disjoint from $I(k, n)$. Also $I(k, n)$ and $I(2nk, n+1)$ have a common centre $k/x_n = 2nk/x_{n+1}$.

Let

$$E_n = \bigcup_{-x_n<k<x_n} I(k, n); \qquad Y_n = E_1 \cap E_2 \cap \cdots \cap E_n.$$

There is clearly a set N_n of 3^{n-1} integers such that

$$Y_n = \bigcup_{k\in N_n} I(k, n).$$

The set $Y = \bigcap_{n=1}^{\infty} Y_n$ is closed and contains the set H of all centres of $I(k, n)$ with $k \in N_n$. Further, each point $y \in Y$ lies in infinitely many of the intervals $I(k, n)$ $(k \in N_n)$ and consequently every neighborhood of y contains points of H. Hence Y is nonempty and perfect and so cannot be enumerable.

If $y \in Y$, $y \neq 0$, then for every integer $n \geqq 1$ there is an integer $k_n \in N_n$ such that $y \in I(k_n, n)$; i.e. $|yx_n - k_n| \leqq 1/4n$ and so

$$x_n \to 0 \pmod{1/y}.$$

It follows that in this case the exceptional set A is not enumerable.

In a note which appeared in this Monthly, (1964, p. 804), after the present article was accepted for publication, Dr. Paul Erdös demonstrated *inter alia* the non-enumerability of the above set A. The proof here given is somewhat simpler than his.

Editorial note. Prof. I. J. Schoenberg has pointed out that in his solution to problem 5090, this Monthly, 71 (1964) p. 333, there is an unfortunate misprint. In line 6 from the bottom of page 333 the limits of summation should read 1, n, not 1, ∞.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

THE IMPOSSIBILITY OF A DIVISION ALGEBRA OF VECTORS IN THREE DIMENSIONAL SPACE

KENNETH O. MAY, Carleton College

When the student discovers that plane vectors, under the identification $(a, b) = a + bi$, satisfy all the field axioms (i.e. are a commutative division algebra over the reals), he ought to wonder whether something similar might be done with vectors in space. When he meets the dot and cross products, he ought to ask why the latter (in spite of its plausible interpretation) does not specialize to the plane and whether something more like the algebra of numbers could be constructed.

The historical record attests to the naturalness and importance of such questions. Indeed they seem to have troubled mathematicians as soon as the nature of complex numbers began to become clear. After many years of pondering imaginaries and of experimenting with operations in space, Gauss in 1831 wrote (in [1]): "The writer has reserved for himself . . . the question why the relations between things that make up a manifold of more than two dimensions cannot provide quantities admissible in universal arithmetic." In 1833 Hamilton (see [2]) arrived at the first really modern treatment of complex numbers as ordered pairs of reals and immediately attempted a generalization to triples. After ten years of intensive and futile experimentation, he took the leap to quaternions without proving the impossibility in 3-space. Grassmann [3] began with efforts to develop an algebra of space vectors and went on to his very general algebras without being able to generalize complex numbers to three dimensions or to prove the impossibility of doing so. De Morgan and others experimented with multiple algebras. Weierstrass in his lectures from 1861 on discussed the Gaussian question ([4] p. 361; [5] p. 312; [6] pp. 24–27), but Hankel in 1867 first printed a proof that no hypercomplex number system could satisfy all the laws of algebra. He wrote ([7] pp. 106–108) " . . . thus is answered the question whose solution was promised but not given by Gauss." Another answer was given after ten years by Frobenius [8], C. S. Peirce [9], E. Cartan [10], and Grisseman [11], who proved that only one more division algebra, namely quaternions, is made possible by dropping the commutativity of multiplication. As recently as 1958, Milnor, Bott, and others [12] proved that the only possible division algebras over the reals (without assuming either the commutative or associative laws of multiplication) are of dimension 1 (the reals), 2 (complex numbers), 4 (quaternions), and 8 (Cayley numbers). The sagacity of Cartan's remark of 1908 ([4] p. 362) that " . . . a definitive answer, if one exists, can only be given by the whole ulterior development of algebra and analysis," may be verified by observing the role played by the Gaussian question in general algebra. (See, for example, [13] Chs. 3–6, [14] Ch. V.)

Although the results cited above make abundantly clear the restrictions on algebra in three dimensions, the literature does not contain an accessible and understandable answer to the question most naturally asked by a student of elementary calculus today, namely, whether multiplication of vectors in 3-space can be so defined that there results a natural generalization of complex numbers just as these were the natural generalization of real numbers. More precisely, we consider the set of triples of reals (x, y, z) with the identifications $(x, 0, 0)=x$ and $(x, y, 0)=(x, y)=x+iy$, with the usual definitions of addition and multiplication by a scalar, and (hopefully) with a multiplication of vectors somehow defined so that all the field properties hold.

That these demands are very inconsistent the student can easily discover for himself if he experiments without getting bogged down (as did Hamilton!) in trying particular definitions of multiplication. For example, letting $(x, y, z)=x+yi+zj$, closure under multiplication implies that for some real a, b, c

$$ij = a + bi + cj. \tag{1}$$

If we multiply both members by i and substitute from (1) in the result, we find

$$(ac - b) + (a + bc)i + (c^2 + 1)j = 0. \tag{2}$$

Hence, $c^2+1=0$ and c is not real. (Compare [**8**], [**9**].)

Since the above argument does not use commutativity, we have proved the impossibility in three dimensions even if that property is not required. Associativity was used when we multiplied (1) by i, but the following example of a zero product neither of whose factors is zero shows that even with neither associativity nor commutativity the other properties of a division algebra cannot hold in three dimensions.

$$(i - c)(ac - b + (bc + a)i + (c^2 + 1)j) = 0. \tag{3}$$

A slightly more complicated example in a more general context was given by Dickson in 1935 ([**15**] pp. 113–115). Others may be found by examining the equation $(A+Bi+Cj)(D+Ei+Fj)=0$.

There are other questions suitable for student investigation. Would a new definition of multiplication of complex numbers make a difference? (See [**4**] pp. 351–353 or [**16**]). What about dimensions higher than three? Complex coefficients? (Consider $j^2=\alpha+\beta i+\gamma j$.) Peirce's proof in [**9**] is readable without extensive background and may lead the student toward deeper study. The Gaussian question, which inspired so much of the development of modern algebra, might be used to good effect as motivation in teaching today.

Written at the University of California, Berkeley, during tenure as an NSF Science Faculty Fellow.

References

1. C. F. Gauss, Review of his *Theoria residuorum biquatraticorum, commentatio secunda*, in *Werke* II, 178. See also *Werke* I, 71; II, 109–110, 174–178; VIII, 353–362 (quaternions before Hamilton!); X:1, 396–397, 436–437; X:2:1, 54–56.

2. W. R. Hamilton, Lectures on Quaternions, 1853, preface. See also C. C. MacDuffee, Algebra's debt to Hamilton, Scripta Math., 10 (1944) 25–33.

3. H. Grassmann, Die lineale Ausdehnungslehre, 1844, preface, or in his *Werke*, I:1, 7–16.

4. E. Cartan and E. Study, Nombres Complexes, in *Encyclopédie des sciences mathématiques*, Paris, 1908, 1:1:3, 329–488.

5. K. Weierstrass, Zur Theorie der aus n Haupteinheiten gebildeten complexen Grössen, *Werke*, II, 311–339.

6. Ernst Kossak, Die Elemente der Arithmetik, Berlin, 1872. First publication of the Weierstrass argument based on his lectures of 1865–1866.

7. H. Hankel, Theorie der complexen Zahlensysteme, Leipzig, 1867.

8. F. G. Frobenius, Über lineare Substitutionen und bilineare Formen, J. Reine Angew. Math., 84 (1878) 1–63.

9. C. S. Peirce, On the algebras in which division is always possible, Amer. J. Math., 4 (1881) 225–229.

10. E. Cartan, Les groupes bilinéaires et les systèmes de nombres complexes, Ann. Fac. Sci. Univ. Toulouse, 12 (1898) B1–B99.

11. F. X. Grisseman, Elementarer Nachweiss des Satzes von Frobenius über die Ausnahmstellung der Quaternionen unter den complexen Zahlensystemen von mehr als zwei Einheiten, Monatsh. Math. und Phys., 11 (1900) 132–147.

12. J. Milnor, Some consequences of a theorem of Bott, Annals of Math., 68 (1958) 444–449.

13. A. A. Albert, ed., Studies in Modern Algebra, MAA Studies in Math., 1963.

14. A. G. Kurosh, Lectures on General Algebra, tr. by K. A. Hirsch, New York, 1963.

15. L. E. Dickson, Linear algebras with associativity not assumed, Duke Math. J., 1 (1935) 113–125.

16. M. L. Smith, A problem in the multiplication of number pairs, Delta-Epsilon, 3 (1963) 49–52.

A NEW MODEL OF THE HYPERBOLIC PLANE

DAVID GANS, New York University

On entering a course in hyperbolic geometry many a student expects to encounter hyperbolas and to have much to do with them. If the course follows a metric rather than projective approach, he will find that its development does not involve hyperbolas directly at all, his closest contact with these curves being via the hyperbolic functions. Nor will he find that there are any hyperbolas in any of the familiar models of the hyperbolic plane. Under these circumstances it is of interest to note that in the model to be described presently hyperbolas play a fundamental role: the branches of a certain set of hyperbolas correspond to the straight lines in hyperbolic plane geometry. As this suggests, the model, unlike familiar models, utilizes the entire Euclidean plane rather than some special part of it.

As a first step in describing the model let us consider the mapping in the Euclidean plane with equations

$$x' = \frac{x}{\sqrt{(x^2 + y^2 + 1)}}, \qquad y' = \frac{y}{\sqrt{(x^2 + y^2 + 1)}}. \tag{1}$$

Its inverse is

$$(2) \qquad x = \frac{x'}{\sqrt{(1 - x'^2 - y'^2)}}, \qquad y = \frac{y'}{\sqrt{(1 - x'^2 - y'^2)}}.$$

The mapping, which we shall denote by T, is topological. It sends the plane into the inside of the unit circle γ with center at O. This occurs radially, points moving toward O along the lines through O so that circles with center O go into smaller such circles. We shall determine the curves which map into the chords of γ.

Taking the diameters of γ first, it is clear from what was said above that each is the image of the straight line through O on which it lies. Next we consider the chord $x' = a$, where $0 < a < 1$. It is the image of the curve

$$(3) \qquad \frac{x}{\sqrt{(x^2 + y^2 + 1)}} = a.$$

Squaring and simplifying this, we get

$$(4) \qquad (1 - a^2)x^2 - a^2y^2 = a^2,$$

an hyperbola with foci on the x-axis, with intercepts $\pm a/\sqrt{(1-a^2)}$ on this axis, and with asymptotes

$$y = \frac{\pm\sqrt{(1 - a^2)}}{a} x.$$

Therefore (3) is the right-hand branch of this hyperbola, and the given chord is its image. Clearly, the chord $x' = -a$ is the image of the left-hand branch. The position of each branch relative to the corresponding chord is particularly simple: the lines joining the endpoints of the chord to O are the asymptotes of the branch.

The remaining chords could be investigated in similar fashion, but an alternative procedure is easier and gives one more insight. It has been shown [1] that T can be obtained as the resultant of two projections. In the first of these we use (0, 0, 1) as center of projection and project the plane onto the lower half of the unit sphere whose center is (0, 0, 1). Then we project this hemisphere, excluding its boundary, back onto the plane orthogonally. The symmetry of these projections is such that the two curves which T maps into any two equal chords are congruent, and each is located in the same way relative to its image as is the other. This is apparent if one starts out with the two equal chords and applies the projections in the reverse order. Since each of the chords that remained to be investigated is equal to one of the chords $x' = a$, where $0 < a < 1$, we see that it, too, is the image of a branch of an hyperbola, and that its position relative to this branch is the same as that of the equal chord $x' = a$ relative to the branch of which it is the image.

We can now state the following two steps for determining the hyperbolic branch corresponding to any chord AB which is at a positive distance k from O: (1) On the ray with endpoint O which is perpendicular to AB take V so that the length of OV is $k/\sqrt{(1-k^2)}$. V is then the vertex of the hyperbolic branch and the latter is symmetrical to line OV. (2) Draw the rays OA, OB. They are the asymptotes of the branch. (Fig. 1 illustrates this. Since we are concerned with only one branch, and not the complete hyperbola, we have taken the asymptotes to be rays rather than complete lines. We shall adhere to this in what follows.)

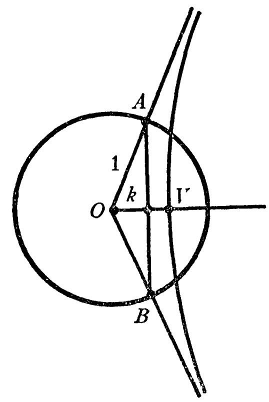


FIG. 1

Thus T maps the entire plane topologically onto the interior of γ so that the chords of γ which are diameters are the images of the straight lines through O, and all other chords are the images of the hyperbolic branches just described. As a variable chord parallel to a diameter approaches the latter, the eccentricity of the corresponding hyperbolic branch becomes infinite and the branch approaches the straight line containing the diameter. In view of this we shall broaden the meaning of the term "hyperbolic branch" so as to have it include each straight line through O. Then we can say simply that the chords of γ are the images of the hyperbolic branches.

It is known that the hyperbolic plane, too, can be mapped topologically onto the interior of γ, this time so that the hyperbolic straight lines correspond to the chords of γ (see [2], Chap. IV, for example). In particular, two chords meeting within γ correspond to two intersecting hyperbolic lines, two chords meeting in a point of γ correspond to two asymptotic parallels, and two chords which do not meet correspond to two nonasymptotic parallels. The resultant of this mapping and T^{-1}, in this order, is therefore a topological mapping of the entire hyperbolic plane onto the entire Euclidean plane in which the hyperbolic branches are the images of the hyperbolic straight lines. In studying this model of the hyperbolic plane one naturally uses the chords of γ as links between the

hyperbolic branches and the hyperbolic straight lines. Thus, two hyperbolic branches will correspond to two intersecting hyperbolic straight lines, two asymptotic parallels, or two nonasymptotic parallels according as, in the mapping T, those branches correspond to chords meeting within γ, meeting on γ, or not meeting at all.

Our mapping of the hyperbolic plane on the Euclidean plane does not go so far as to preserve distance. Nevertheless, certain metric properties of the hyperbolic branches, in terms of Euclidean distance, do correspond to metric properties of hyperbolic straight lines, in terms of hyperbolic distance. We shall mention five such properties.

(1) Each hyperbolic branch resembles a hyperbolic straight line in being infinitely long.

(2) Two hyperbolic branches which meet in a point diverge more and more as they recede from the point in either direction, just as do the two intersecting hyperbolic straight lines to which they correspond. This can be verified by means of Fig. 2, which shows two hyperbolic branches meeting in P, the point $P' = T(P)$, the chords AB, CD to which these branches correspond, and the four distinct asymptotes OA, OB, OC, OD.

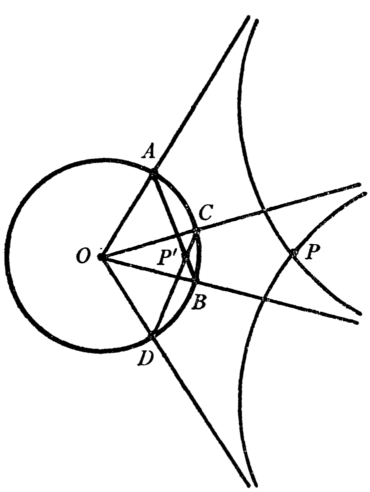


FIG. 2

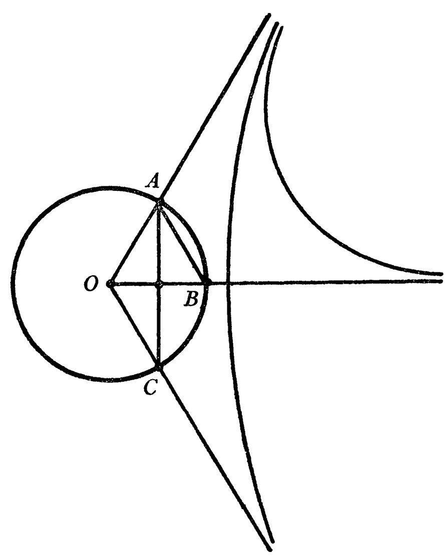


FIG. 3

(3) Two hyperbolic branches which correspond to two asymptotic parallels are themselves asymptotic, converging in one direction and diverging in the other. Figures readily show why this must be so. Fig. 3, for example, shows two such branches and the chords AB, AC to which they are related by T. Since the chords meet on γ, the branches have a common asymptote, which gives the direction in which they converge. The other asymptotes being distinct, the branches diverge in approaching them.

(4) Through a point outside of a hyperbolic branch there pass exactly two hyperbolic branches which are asymptotic to it, and they are asymptotic to it in opposite directions. For T maps the given point and branch into a point P' and chord AB within γ (P' not on AB), and there are just two chords through P' which meet AB on γ (Fig. 4).

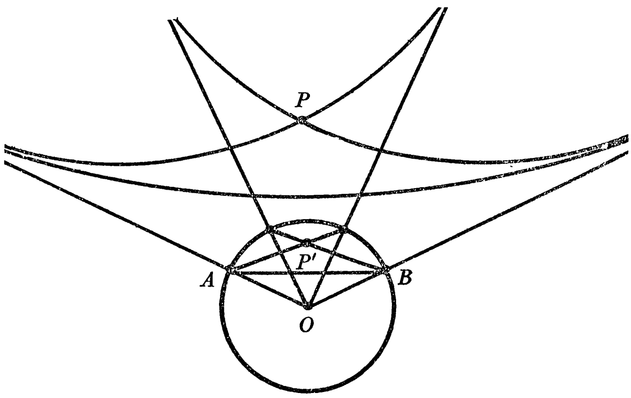


FIG. 4

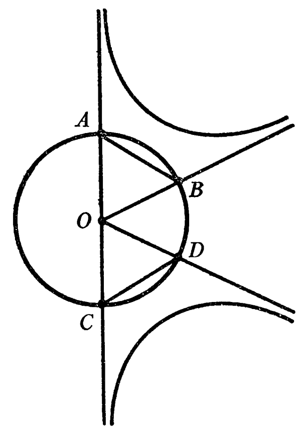


FIG. 5

(5) Two hyperbolic branches which correspond to two nonasymptotic parallels diverge from one another in all directions, as do the two parallels. This follows from the fact that the chords AB, CD which correspond to the two branches do not meet, and hence that the asymptotes OA, OB of one branch are distinct from the asymptotes OC, OD of the other (Fig. 5).

References

1. D. Gans, Bounded models of the Euclidean plane, this MONTHLY, 51 (1954) 26–30.
2. R. Baldus, Nichteuklidische Geometrie, de Gruyter, Berlin, 1927.

INTERVAL INDUCTION

GERALD JUNGCK, Bradley University

Both Ford [1] and Duren [2] have presented methods for proving statements concerning closed intervals of real numbers. The purpose of this paper is to offer another such method. The method is similar in form to that of finite induction—hence the title, "Interval Induction." We shall first state and prove the induction principle, Theorem A. We then illustrate the application of this theorem by proving three well-known theorems.

THEOREM A. *Let $P(C)$ denote a proposition pertaining to a nonempty subset C of $[a, b]$. Then $P([a, b])$ is true provided the following conditions are satisfied.*

(1) *$P(\{a\})$ is true.*

(2) *If $c\in[a, b]$, and if for any $\epsilon>0$ there is a point x in $(c-\epsilon, c]$ such that $P([a, x])$ is true, then there is a $\delta>0$ such that $P([a, c+\delta]\cap[a, b])$ is true.*

Proof. Let $S=\{x\in[a, b] \mid P([a, x])$ is true$\}$. By (1), $P([a, a])$ is true, so that $S\neq\emptyset$. Moreover, S is bounded above by b and hence has a l.u.b., c, in $[a, b]$. But then, for any $\epsilon>0$, $S\cap(c-\epsilon, c]\neq\emptyset$ and hence there is a point x in $(c-\epsilon, c]$ such that $P([a, x])$ is true. By (2), there is a $\delta>0$ such that $P([a, c+\delta]\cap[a, b])$ is true. If $c<b$ and $d=\min(c+\delta, b)$, then $P([a, d])$ is true. This is impossible since $d>c$ and c is the l.u.b. of S. It follows that $c=b$, and that $P([a, b])$ is true.

The first example is a proof of the Bolzano Weierstrass theorem for the real numbers. In this proof we employ the fact that, given the ordinary topology for R (the real numbers), a point p in R is *not* a limit point of a set E in R if and only if there exists an $\epsilon>0$ such that $[p-\epsilon, p+\epsilon]\cap E$ is finite.

THEOREM. *Every bounded infinite subset E of R has a limit point in R.*

Proof. Suppose that E has no limit point in R. Since E is bounded, there is an interval $[a, b]$ such that $E\subset[a, b]$. Let $P(C)$ denote the proposition, "$E\cap C$ is finite." Now $P(\{a\})$ is trivially true, and (1) of Theorem A is satisfied. Let c be any point of $[a, b]$. By assumption, c is not a limit point of E so that there is an $\epsilon>0$ such that $[c-\epsilon, c+\epsilon]\cap E$ is finite. But, if c satisfies the hypothesis of (2) in Theorem A, there is a point x in $(c-\epsilon, c]$ such that $P([a, x])$ is true; i.e., $[a, x]\cap E$ is finite. Since $[c-\epsilon, c+\epsilon]\cap E$ is finite, $[a, c+\epsilon]\cap E$ is finite, and $P([a, c+\epsilon]\cap[a, b])$ is true. Thus (2) of Theorem A is satisfied so that $P([a, b])$ is true, i.e., $E\cap[a, b]$ is finite. But this is impossible since $E\cap[a, b]=E$, and E is infinite. It follows that E has a limit point in R.

We next prove, without appeal to the boundedness of $f(x)$, that a real-valued function f continuous on a closed interval attains a maximum value.

THEOREM. *If f is continuous on $[a, b]$, $f(x)$ attains a maximum value on $[a, b]$.*

Proof. Suppose that $f(x)$ does not attain a maximum value on $[a, b]$. Let $P(C)$ denote the proposition, "There is a point y in $[a, b]$ such that $f(y)>f(x)$ for each $x\in C$." Since $f(a)$ is not maximal, $P(\{a\})$ is true, and (1) of Theorem A is satisfied.

Let c be any point in $[a, b]$ such that for any $\epsilon>0$ there is an $\bar{x}\in(c-\epsilon, c]$ for which $P([a, \bar{x}])$ is true. Since $f(c)$ is not maximal, we can choose a point y in $[a, b]$ such that $f(y)>f(c)$, and the continuity of f permits an $\epsilon>0$ such that $f(y)>f(x)$ for all

$$x\in[c-\epsilon, c+\epsilon]\cap[a, b].$$

Hence there is a point $\bar{x}$ in $(c-\epsilon, c]$ and a point y' in $[a, b]$ such that $f(y')>f(x)$ for all x in $[a, \bar{x}]$. Thus, if $f(d)=\max(f(y), f(y'))$, we have $f(d)>f(x)$ for all

$$x\in[a, c+\epsilon]\cap[a, b].$$

This implies that $P([a, c+\epsilon]\cap[a, b])$ is true, so that (2) of Theorem A is satisfied. By Theorem A, $P([a, b])$ is true. But $P([a, b])$ yields the embarrassment: $f(y)>f(y)$ for some point y in $[a, b]$.

In our third and last example we employ the fact that a function continuous

on a closed interval is bounded (this follows easily from Theorem A) to prove that the image of a closed interval under a continuous mapping is a closed interval. We note that this last result is obtained without the traditional two-phase approach of first proving extreme value attainment, and then proving the Intermediate Value Theorem.

THEOREM. *Let f be continuous on $[a, b]$ and let M and m denote the l.u.b. and the g.l.b. respectively, of the range of f. If K is any number such that $m \leqq K \leqq M$, there exists an x in $[a, b]$ such that $f(x) = K$.*

Proof. Suppose there is no $x \in [a, b]$ such that $f(x) = K$. Then $f(a) \neq K$, and we assume that $f(a) < K$. (The other possibility, $f(a) > K$, merely reverses the inequalities involving $f(x)$ in the following argument.) Let $P(C)$ denote the proposition, "There is an $\epsilon > 0$ such that $f(x) < K - \epsilon$ for all x in C." It is immediate that $P(\{a\})$ is true.

Let $c \in [a, b]$ and suppose that for any $\delta > 0$ there is an $\bar{x}$ in $(c - \delta, c]$ such that $P([a, \bar{x}])$ is true. By assumption, $f(c) \neq K$. Also, $f(c) \not> K$, since, otherwise, continuity permits a $\delta > 0$ such that $f(x) > K$ for all x in

$$[c - \delta, c + \delta] \cap [a, b].$$

Hence $f(c) < K$, and, by continuity, there is a positive δ such that $f(x) < K - \epsilon(c)$ for all $x \in [c - \delta, c + \delta] \cap [a, b]$, where $\epsilon(c) = (K - f(c))/2$. But then we can pick an $\bar{x} \in (c - \delta, c]$ and an $\epsilon > 0$ such that $f(x) < K - \epsilon$ for all $x \in [a, \bar{x}]$. Thus, if $\epsilon' = \min(\epsilon, \epsilon(c))$, then $f(x) < K - \epsilon'$ for all x in $[a, c + \delta] \cap [a, b]$, and $P([a, c + \delta] \cap [a, b])$ is true.

We now have the anticipated contradiction. For by Theorem A, $P([a, b])$ is true. But $P([a, b])$ reads, "There is an $\epsilon > 0$ such that $f(x) < K - \epsilon < K \leqq M$ for all $x \in [a, b]$." This is impossible since M is the l.u.b. of the range of f.

References

1. L. R. Ford, Interval additive propositions, this MONTHLY, 64 (1957) 107–109.
2. W. L. Duren, Mathematical induction in sets, this MONTHLY, 64, October, part 2, (1957) 19–21.

A SIMPLIFIED PROOF OF A SUFFICIENT CONDITION FOR A POSITIVE DEFINITE QUADRATIC FORM

S. M. SAMUELS, Purdue University

We present an inductive proof of the statement that positivity of the leading principal minors of the matrix of a quadratic form implies positive definiteness of the form. The chief interest of the proof is its simplicity. It compares favorably with the proof given by C. J. Seelye [1], which, in turn was offered in part as a simplification of the proof given in most textbooks.

The proof is an immediate consequence of the following simple

LEMMA. *Let*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{a}_n \\ \mathbf{a}_n' & a_{nn} \end{bmatrix}$$

be an n-th order symmetric matrix, with $\mathbf{A}_{n-1}$ *the (symmetric) submatrix consisting of the first* $n-1$ *rows and columns of* $\mathbf{A}$. *Then, if* $\mathbf{A}_{n-1}$ *is nonsingular, there is a nonsingular matrix P such that*

$$\mathbf{P}'\mathbf{A}\mathbf{P} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{0} \\ \mathbf{0}' & b_{nn} \end{bmatrix}.$$

If, in addition, $\det \mathbf{A}_{n-1}$ *and* $\det \mathbf{A}$ *are positive, then* b_{nn} *is positive.*

Proof. Since $\mathbf{A}_{n-1}$ is nonsingular, $\mathbf{a}_n$ has a unique representation as a linear combination of the columns of $\mathbf{A}_{n-1}$. Let $\alpha_1, \cdots, \alpha_{n-1}$ be the coefficients of the corresponding columns of $\mathbf{A}_{n-1}$ in this representation. Let $\mathbf{P}'$ be the matrix obtained from the identity matrix by replacing its last row by

$$[-\alpha_1, \cdots, -\alpha_{n-1}, 1].$$

Then $\mathbf{P}$ is nonsingular and, by the symmetry of $\mathbf{A}$, is the desired matrix.

Since

$$\begin{aligned} \det \mathbf{P}'\mathbf{A}\mathbf{P} &= (\det \mathbf{P})^2 \det \mathbf{A} \\ &= b_{nn} \det \mathbf{A}_{n-1}, \end{aligned}$$

the second part of the lemma follows immediately.

We now prove the desired theorem, assuming, without loss of generality, that the matrix of the quadratic form is symmetric.

THEOREM. *If the leading principal minors of the n-th order symmetric matrix* $\mathbf{A}$ *are all positive, then the quadratic form* $\mathbf{x}'\mathbf{A}\mathbf{x}$ *is positive definite.*

Proof. The theorem is obviously true for $n=1$. Assume it is true for $n-1$ and let $\mathbf{A}$ be an nth order symmetric matrix with positive leading principal minors. Then $\mathbf{A}$ satisfies the hypotheses of the lemma. Now positive definiteness of $\mathbf{x}'\mathbf{A}\mathbf{x}$ is equivalent to positive definiteness of

$$\mathbf{x}'\mathbf{P}'\mathbf{A}\mathbf{P}\mathbf{x} = \mathbf{x}_{n-1}'\mathbf{A}_{n-1}\mathbf{x}_{n-1} + x_n^2 b_{nn}$$

where $\mathbf{x}_{n-1}$ denotes the first $n-1$ variables. The form is positive definite since the first term is positive by the inductive hypothesis (every leading principal minor of $\mathbf{A}_{n-1}$ is a leading principal minor of $\mathbf{A}$) and the second term is positive by the lemma.

Reference

1. C. J. Seelye, Conditions for a positive-definite quadratic form established by induction, this MONTHLY, 65 (1958) 355–356.

A NOTE ON THE VOLUME OF A SIMPLEX

P. STEIN, London

1. Let X_r, $r=1, 2, \cdots, n+1$, be $n+1$ points in an n-dimensional space. Suppose $X_r=(x_{r1}, x_{r2}, \cdots, x_{rn})$. The simplex $V=V(X)$ spanned by these points is the set of points

$$X = \sum_{r=1}^{n+1} \alpha_r X_r \qquad \alpha_r \geqq 0,\ r = 1, 2, \cdots, n+1,\ \sum_{r=1}^{n+1} \alpha_r = 1.$$

We shall give a simple proof of the known fact that the volume $|V(X)|$ of this simplex is given by

$$|V(X)| = \pm \frac{1}{n!} \det \begin{bmatrix} X_1, \cdots, X_{n+1} \\ 1, \cdots, 1 \end{bmatrix},$$

where X_r represents the column vector $(x_{r1}, \cdots, x_{rn})'$.

2. Consider an affine transformation $X=AY+B$, $\det A\neq 0$, where B is an n-dimensional column vector.

Then $[X_2-X_1, \cdots, X_{n+1}-X_1]=A[Y_2-Y_1, \cdots, Y_{n+1}-Y_1]$. We have

$$\det \begin{bmatrix} X_1, \cdots, X_{n+1} \\ 1, \cdots, 1 \end{bmatrix} = (-1)^{n+1} \det [X_2 - X_1, \cdots, X_{n+1} - X_1]$$

and

$$\det \begin{bmatrix} Y_1, \cdots, Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix} = (-1)^{n+1} \det [Y_2 - Y_1, \cdots, Y_{n+1} - Y_1].$$

It follows that

$$\det \begin{bmatrix} X_1, \cdots, X_{n+1} \\ 1, \cdots, 1 \end{bmatrix} = \det A \det \begin{bmatrix} Y_1, \cdots, Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix}. \tag{1}$$

We note two special cases of (1):
(1a) A orthogonal: in this case, $\det A = \pm 1$,
(1b) A scalar, say $A=\rho I$, $\rho\neq 0$; in this case, $\det A=\rho^n$.

3. We now suppose that the simplex $V(X)$ is nondegenerate, i.e., the points $X_1, \cdots, X_{n+1}$ span the whole n-dimensional space and the determinant on the left of (1) does not vanish. In this case simple geometric considerations establish the existence of an orthogonal matrix A and a vector B such that for the transformation $X=AY+B$,

$$Y_1 = (y_0, 0, \cdots, 0), \quad Y_r = (0,\ y_{r2}, \cdots, y_{rn}), \quad r > 1.$$

In virtue of (1a) it follows that $|V(Y)| = \pm|V(X)|$. Now the volume of $V(Y)$ can be written as

$$|V(Y)| = \int dy_1 dy_2 dy_3 \cdots dy_n,$$

where the n-tuple integral is evaluated over all Y such that

$$Y = \alpha_1 Y_1 + \sum_{r=2}^{n+1} \alpha_r Y_r,$$

where $\alpha_r \geqq 0$, $r=1, 2, \cdots, n+1$ and $\sum_{r=2}^{n+1} \alpha_r = 1-\alpha_1$. It is clear that

$$(2) \qquad |V(Y)| = \int_0^{y_0} dy_1 \left[\int dy_2 \cdots dy_n \right].$$

4. We now use induction. The required formula is true for the 2-dimensional case. Assume it is true for the $(n-1)$-dimensional case. Then for the simplex spanned by $z_2, \cdots, z_{n+1}$ we have

$$V(Z) = \int dz_2 \cdots dz_{n+1} = \pm \frac{1}{(n-1)!} \det \begin{bmatrix} Z_2, \cdots, Z_{n+1} \\ 1, \cdots, 1 \end{bmatrix},$$

where the integral is evaluated over all $z_2, \cdots, z_{n+1}$ for which

$$Z = \sum_{r=2}^{n+1} \beta_r Z_r \qquad \beta_r \geqq 0, \ \sum_{r=2}^{n+1} \beta_r = 1.$$

Now if we write

$$\begin{aligned} Z_r &= (1-\alpha_1) Y_r \\ \beta_r &= \alpha_r (1-\alpha_1)^{-1} \end{aligned} \qquad r = 2, \cdots, n$$

then it is clear that the inner integral in (2) is

$$\pm \frac{1}{(n-1)!} \det \begin{bmatrix} (1-\alpha_1) Y_2, \cdots, (1-\alpha_1) Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix}$$
$$= \pm (1-\alpha_1)^{n-1} \frac{1}{(n-1)!} \det \begin{bmatrix} Y_2, \cdots, Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix}.$$

Hence

$$|V(Y)| = \pm \frac{1}{(n-1)!} \det \begin{bmatrix} Y_2, \cdots, Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix} \int_0^{y_0} (1-\alpha_1)^{n-1} dy_1.$$

Since $y_1 = \alpha_1 y_0$, $0 \leqq \alpha_1 \leqq 1$, we have

$$\int_0^{y_0} (1-\alpha_1)^{n-1} dy_1 = \int_0^1 (1-\alpha_1)^{n-1} y_0 \, d\alpha_1 = y_0/n,$$

and so

$$|V(y)| = \pm \frac{1}{n!} y_0 \det \begin{bmatrix} Y_2, \cdots, Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix}$$

$$= \pm \frac{1}{n!} \det \begin{bmatrix} Y_1, \cdots, Y_{n+1} \\ 1, \cdots, 1 \end{bmatrix}.$$

This completes the proof.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College, and
JOHN A. BROWN, University of Delaware

All material for this department should be sent to John R. Mayor, 1515 Massachusetts Avenue, N.W., Washington, D. C. 20005.

MADISON PROJECT ACTIVITIES FOR 1965–1966: REPORT ON UNFINISHED BUSINESS

R. B. DAVIS, Syracuse University and Webster College

The "modern" mathematics curriculum efforts have been lately receiving notice in newspaper comic strips and in the songs of Tom Lehrer, and may consequently appear to be in an advanced stage of completion, if not obsolescence. In our view this appearance is deceptive: the "mathematics revolution" in pre-college school programs has not taken place. Advances have clearly been made in some areas and in some respects, but what has been accomplished is far less than what is needed or what is actually possible. The present note reports primarily unfinished business which may be of interest to mathematicians, and which will depend upon the participation of mathematicians if such work is to be carried forward.

By way of background, the Madison Project of Syracuse University and Webster College is a curriculum evolution project, supported mainly by the National Science Foundation and by the Bureau of Research of the United States Office of Education, that has been active since 1957, and has been concerned primarily with programs of study in mathematics for nursery school through high school, plus undergraduate college education of prospective teachers, and in-service study by presently active teachers. Some Project materials combine science with mathematics.

1. An "Inside-out" approach to teacher education. Most teacher education in mathematics (where any such thing exists at all) uses the format of teaching mathematics to the teachers, and then leaving the teachers with the task of devising appropriate classroom experiences for children. With in-service elementary teachers this approach is not usually successful; the mathematical ideas are in

dents, not by the faculty. The course is taught jointly by Professor William Walton of the Physics Department, and by Professor Robert B. Davis of the Mathematics Department.

References

1. Robert B. Davis, The Madison Project—A Brief Introduction to Materials and Activities. The Madison Project, 1962 (revised 1965).

2. Robert B. Davis, A Modern Mathematics Program as it Pertains to the Interrelationship of Mathematical Content, Teaching Methods and Classroom Atmosphere. The Madison Project, 1965. Report submitted to the Commissioner of Education, U. S. Department of Health, Education and Welfare.

STUDY OF SCIENCE AND MATHEMATICS PROGRAMS IN THE ELEMENTARY SCHOOL

J. R. MAYOR AND H. H. WALBESSER, JR., AAAS and the University of Maryland

Background. The Commission on Science Education of the American Association for the Advancement of Science, supported by the National Science Foundation since 1962, has developed an experimental program in elementary school science, known as *Science—A Process Approach* (*S—APA*). The materials consist of teacher texts for a sequential science program for kindergarten through grade 6. The program is different from others in that there are teacher texts but no pupil texts, the primary goal is to develop children's competence in the processes of science, there is built into each exercise a design for evaluation, and the program includes some experiences on topics usually considered a part of the elementary school mathematics program.

In the third experimental edition (published, fall 1965) of *S—APA* exercises on mathematical topics have been included at each grade level, K through 6. These exercises, in no sense or at no grade level constitute a complete or sequential program in mathematics. They are intended to supplement the regular mathematics program of the school and are introduced as preparation for a particular science exercise or exercises to follow. Usually the introduction of an exercise on a mathematical topic is made when the probability is high that the mathematical competence needed for an ensuing science exercise will not yet have been acquired in the school's regular mathematics program, or when it is believed that special aspects of a mathematical skill, already introduced in the mathematics program, should be given a particular kind of emphasis in preparation for a subsequent science experience.

In the preparation of the exercises of *S—APA*, the scientists and teachers were asked not to delay the teaching of a topic or process in science just because the students might not possess the mathematical competence needed for it. The writers were assured that the necessary mathematical experiences would be introduced as needed. In addition, those responsible for writing the mathe-

matical exercises were asked so far as possible to introduce the mathematical topic in a science setting. The best examples of the response of the writers to this request are in the exercises on graphing and an exercise with the title, Dividing to Find Rates and Means, for grade 2.

The exercises on mathematical concepts are an integral part of the entire hierarchy of competencies upon which *S—APA* is based. One of the eight basic processes of the primary grade sequence is the process of Using Numbers. Some mathematical experiences appear in sequences for other processes such as Using Space/Time Relations, Measuring, Classifying, and Communicating. Competencies, usually classified as mathematical, play an important role in the intermediate grade materials particularly in the sequence for the integrated process, Interpreting Data.

The Study. In the summers of 1963 and 1964, 8-week writing sessions for the development of *S—APA* were held at Stanford University in close proximity to the writing sessions of the School Mathematics Study Group. There was frequent opportunity for communication between members of the two writing teams. In the spring of 1965 representatives of SMSG and of the Commission on Science Education met to consider issues of mutual concern. From this meeting, plans were developed for identification of a number of school systems in which some teachers would teach both SMSG mathematics and *S—APA* during this school year.

The staff of the Commission on Science Education has developed the Science-Mathematics Study described in this report. The staff of the School Mathematics Study Group is making a separate and independent investigation of the class activities in schools using both programs.

In the remainder of this report the term "experimental mathematics" program is used to avoid the ambiguity of the terms "new mathematics" or "modern mathematics." There is no intent to imply that the elementary mathematics program of SMSG is or is not still in the experimental stage.

Purpose of the Science-Mathematics Study. The study is planned to provide information which may be useful in resolving a number of important issues related to the teaching of science and mathematics in the elementary school. A number of these issues are suggested by the following questions:

1. What are the effects on school atmosphere, teachers, and pupils of the teaching of two experimental programs (in this instance, science and mathematics) at the same time?

2. In what ways does the study of an experimental mathematics program appear to strengthen or interfere with a program based on *S—APA* and, the other way around, in what ways does the study of the experimental science sequence appear to strengthen or interfere with an experimental mathematics program?

3. Scientists often report that even the best students in mathematics are unable to apply their mathematical competencies in simple applications in

science. Does the concurrent study of two experimental programs (one in mathematics and one in science) make this generalization less tenable?

4. Would changes in the ordering of topics in an experimental mathematics program be beneficial to the experimental science sequence, and particularly to the development of an ability to use mathematics in science experiences?

This year an attempt will be made to obtain research data on growth in mathematical competence (the pre- and post-test) and one answer to question 4. The study will depend upon observations of teachers, administrators, and the Commission staff for information which may suggest answers to the other questions.

Participating Schools. Participation in the science-mathematics study implies the teaching of *S—APA* and the School Mathematics Study Group elementary school materials in at least one class in each grade, kindergarten through grade 6. An exception is made in the case of Lakewood, Ohio schools, in which a mathematics program, other than SMSG is being used.

A number of the schools have several classes at each grade level. In all cases the same teacher is teaching both science and mathematics except in Delaware in grades 5 and 6 and the University of Chicago Laboratory Schools in grades 4, 5, and 6. Special teachers of science and of mathematics are used in these systems.

The participating school systems are: Several districts around Newark, Delaware; Overland Park, Kansas; Campbellsville, Kentucky; Wenatchee, Washington; University of Chicago, Laboratory Schools; Lakewood, Ohio.

TERM PROJECTS IN MATHEMATICS FOR LIBERAL ARTS STUDENTS

H. F. RAHMLOW, Kansas Wesleyan University

Professor Alder has recently written in the MONTHLY describing a suitable course in mathematics for liberal arts students (H. L. Alder, *Mathematics For Liberal Arts Students*, this MONTHLY, 72 (1965) 60–66). For the past year I have been teaching such a course to liberal arts students with wide ranges in ability. I have used a device not mentioned by Professor Alder which seems to have contributed substantially to the success of these courses. This device is the requirement of a term paper or project of each student. Each student is encouraged to select a topic within his major field of interest. Initially most of the students need quite a bit of guidance in selecting a topic, but after a personal conference or two a worthwhile area is readily selected. Such a personal conference also provides the instructor with an excellent opportunity for discussing mathematics with nonscience majors. These discussions in themselves can be of great benefit to the instructor as well as to the student.

It is my opinion and the opinion of the great majority of the students who have taken these courses that there are three main reasons for the success of term papers in a liberal arts mathematics course. Most important, the liberal arts student is able to see, often for the first time, that mathematics permeates

his major field of interest. The student can study worthwhile mathematical ideas from the secure viewpoint of his prime area of competence.

Second, the student is able to see for himself that mathematics is a dynamic field of study and not a dead collection of algorithms. As the student seeks appropriate material for his paper he is forced to confront recent journal and magazine articles as well as text books. Some students are led to discussions with practicing mathematicians, teachers, and others who routinely deal with emerging mathematics. Thus, students can begin to see mathematics as it really is.

Finally, the above two points, finding mathematics in one's own field of interest and seeing mathematics as a dynamic field of study, combine to stimulate interest in the normal course and text material.

A few examples of papers done by my students will serve to emphasize the points made above. An English major investigated the life of Lewis Carroll and the mathematical concepts embodied in his writings. A music major, stimulated by a problem from the text, went on to investigate some of the numerous applications of mathematics to music. Several students who were interested in teaching investigated from a mathematical viewpoint the newer trends in school mathematics and the philosophy behind these trends. All of these students found and studied mathematics in areas where many did not previously know that mathematics existed.

In conclusion, it is my opinion that carefully selected term projects can greatly enrich a course in mathematics for liberal arts students and at the same time substantially increase the students' appreciation for mathematics.

MODERN METHODS OF TEACHING MATHEMATICS TO CULTURALLY DEPRIVED COLLEGE STUDENTS

R. G. PETTIS, Winston-Salem State College

"Modern Mathematics" for college students has not allowed for the formidable deficiencies of the so called "culturally deprived." The barrage of ideas and concepts thrust upon these students as they enter college is too often totally incomprehensible. An impoverished mathematical vocabulary and lack of skills in interpretive reading and critical thinking cause frustration and repeated failure in basic college mathematics courses. My initiation of a research project on methods of teaching modern mathematics to the culturally deprived extended from this premise.

Students for the project were selected from the freshman class. Two classes were organized, both composed of students who had scored from 30–90 percentiles on the California Mathematics Test. With the control group I used suggested procedures and materials outlined in the textbook adopted by the Department of Mathematics.

The concepts which the author presented were too often superfluous and did not contribute to clarifying relationships and continuity of ideas. Subse-

quent tests revealed that a large number of students who had scored in the first quartile on the entrance examination fell to the second and third quartiles.

No textbook was used for the experimental group. Subject matter was selected from a variety of sources. I presented mathematical concepts which emphasized structural development and continuity of ideas. Such selectivity permitted me to use simplified symbolism and to introduce those concepts which I felt were necessary and efficacious for continual development. For example, the set concept has extensive application in the development of the number system, concept of counting, and proof of the properties of respective number systems, and other closely related ideas; I saw no reason, however, to dwell on the tenuous concepts of sets, hence reducing the time spent on the paramount material.

A special effort was made to establish and maintain a practical articulation between the mathematical concepts, interpretive reading ability, and critical and creative thinking. The students of the experimental group were ardently responsive and subsequent tests revealed improvement in reading ability, ability to do creative thinking, and over-all achievement.

My experience with the groups indicated to me the following important conclusions regarding material and methods of teaching modern mathematics:

1. The complicated symbolism and superfluous material should be modified and de-emphasized.

2. More emphasis should be placed on structural development, and continuity of ideas expressed in language which takes into account both limited vocabulary and limited experience.

3. Too much rigor causes students, who are not strongly mathematically inclined, to lose interest, which in turn hampers their achievement.

I am convinced that the culturally deprived are able to profit from modern mathematics. The difficulty is one that can and must be transcended.

REPORTS OF COLLEGE COMMISSIONS

The December, 1965 issue of SCIENCE EDUCATION NEWS contains brief reports of education activities in 1965 of groups of college and university scientists in the various scientific disciplines. Included are reports from the Committee on the Undergraduate Program in Mathematics, the Advisory Council on College Chemistry, the Commission on College Physics, the Committee on Undergraduate Education in the Biological Sciences, the Commission on Engineering Education, the Commission on College Geography, the Committee on Educational Policy in Agriculture, and the Geological Orientation Study. These agencies are supported by the National Science Foundation. Copies of this issue of SCIENCE EDUCATION NEWS are available upon request.

SCIENCE EDUCATION NEWS is published quarterly by the American Association for the Advancement of Science. It is prepared for scientists in colleges and universities, government, and industry and contains reviews of activities in science education. Readers of the MONTHLY may be added to the mailing list upon request. There is no charge.

PROBLEMS AND SOLUTIONS

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, Bloomfield College, Bloomfield, N. J. 07003. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to A. E. Livingston, University of Alberta, Edmonton, Alberta, Canada. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before July 31, 1966.

E 1865. *Proposed by E. O. Thorp, New Mexico State University*

(a) How many distinct hands (i.e. of different values) are there in 5-card draw poker? More precisely, the hands can be grouped into equivalence classes such that the hands in any equivalence class are equal (tie), and such that, given any pair of equivalence classes, all the hands in one equivalence class beat all the hands in the other. The problem is to find how many equivalence classes there are.

(b) Which equivalence class (or classes) determines the median hand?

E 1866. *Proposed by Michael Fried, University of Michigan*

Given n points in the Euclidean plane. Find a ruler and compass construction for the smallest circle containing these n points.

E 1867. *Proposed by R. E. Shafer, Lawrence Radiation Laboratory, University of California*

Show that for $x>0$

$$\tan^{-1} x > \frac{3x}{1 + 2\sqrt{(1 + x^2)}} .$$

E 1868. *Proposed by K. E. Eldridge, University of Colorado*

Let p be any prime. Let t and s be integers such that $t>0$ and $0<s<p^t$. Define integer i as the exponent of the highest power of p dividing s, and integer a as the exponent of the highest power of p dividing the binomial coefficient $\binom{p^t}{s}$. Prove that $a=t-i$.

E 1869. *Proposed by J. A. Burslem, St. Louis University*

(a) Let L be a finite set of pairwise intersecting lines in the plane, and let P be the set of points determined by L. Define the multiplicity of a point $p \in P$

as the number of lines of L passing through p. For each integer $k \geqq 2$, given i_k, the number of points of multiplicity k, determine the number N of regions into which L divides the plane.

(b) Do the analogous problem resulting when "great circle" and "sphere" are substituted for "line" and "plane," respectively.

E 1870. *Proposed by Arun Sanyal, Indian Institute of Technology, Kharagpur*

(1) Prove that the third order determinant (a_{ij}) vanishes identically, where

$$a_{11} = (A + B)DE - (D + E)AB,\ a_{12} = AB - DE,\ a_{13} = A + B - D - E,$$
$$a_{21} = (B + C)EF - (E + F)BC,\ a_{22} = BC - EF,\ a_{23} = B + C - E - F,$$
$$a_{31} = (C + D)FA - (F + A)CD,\ a_{32} = CD - FA,\ a_{33} = C + D - F - A.$$

(2) The same, where

$$a_{11}=\sin(A+B)\cos(D-E)-\sin(D+E)\cos(A-B),$$
$$a_{12}=\cos(A+B)\cos(D-E)-\cos(D+E)\cos(A-B),$$
$$a_{13}=\cos(A+B+D+E),\ a_{23}=\cos(B+C+E+F),\ a_{33}=\cos(C+D+E+A),$$
$$a_{21}=\sin(B+C)\cos(E-F)-\sin(E+F)\cos(B-C),$$
$$a_{22}=\cos(B+C)\cos(E-F)-\cos(E+F)\cos(B-C),$$
$$a_{31}=\sin(C+D)\cos(F-A)-\sin(F+A)\cos(C-D),$$
$$a_{32}=\cos(C+D)\cos(F-A)-\cos(F+A)\cos(C-D).$$

E 1871. *Proposed by R. A. Jacobson, South Dakota State University*

Let $A = \{x \mid x \in [0, 1]$, and the infinite decimal expansion of x consists of not more than 9 distinct integers$\}$. Is A countable or uncountable? Find the measure of A.

E 1872. *Proposed by R. A. Jacobson, South Dakota State University*

(a) Let $U = \{x \mid x \in [0, 1]$, and at least two consecutive digits in the decimal expansion of x are identical$\}$. Find measure U.

(b) Let $V = \{x \mid x \in [0, 1]$, and at least three consecutive digits in the decimal expansion of x are identical$\}$. Find measure V.

(c) Generalize to an arbitrary base b and to k consecutive digits identical.

E 1873. *Proposed by G. A. Heuer and Bruce Erickson, Concordia College*

Let S be a semigroup with cancellation. Assume that for every element $a \in S$, there is an integer $n_a > 1$ such that $a^{n_a} = a$. Is S necessarily a group?

E 1874. *Proposed by W. A. McWorter, University of British Columbia*

Let G be a group of order n^2 and let H be a subgroup of order n. Prove that for any $x \in G$ we have $H \cap x^{-1}Hx \neq 1$.

SOLUTION OF ELEMENTARY PROBLEM

Arithmetic Progression of Integers Prime to n

E 1730 [1964, 912; 1965, 907]. *Proposed by Robert Spira, Duke University*
Let $(a, b)=1$. If n is any integer, there is an x such that $(ax+b, n)=1$.

Editorial Note. Solution I, as printed, is invalid—in fact, there seems to be no way in which it can be corrected without complete revision. D. M. Crystal provides the following counterexamples.

If $(a, b)=1$ and $(b, n)>1$, it does not follow that $(a, n)=1$: take $(a, b)=1$ and $n=aby$, where y is any integer and $a>1$, $b>1$; then $b\mid(b, n)$, $a\mid(a, n)$, implying that $(b, n)>1$ and $(a, n)>1$.

If $(b, n)>1$, $(a, b)=1$, it does not follow that $(a(n+1)+b, n)=1$: take a, b both odd and $n=2b$; then $a(n+1)+b$ and n are both even.

The error was also reported by Harley Flanders.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions to Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before September 30, 1966.

5324 [1965, 915]. *Proposed by Alan Sutcliffe, Cheshire, England*

CORRECTION: The integer in (a) should be 3367.

5370. *Proposed by F. S. Cater, University of Oregon*

Let X be an infinite set with $\alpha=\text{card } X$ such that either $\alpha=\aleph_0$ or α has an immediate predecessor. Prove that there exists a family M of subsets of X such that card $M>\alpha$ and

(1) card $(A_1\cap A_2)<\alpha$ for any two distinct members A_1, A_2 of M,
(2) for any subfamily N of M with card $N=\alpha$, there exists an $A\in M-N$ such that card $[(\bigcup_{B\in M} B)\cap A]=\alpha$.

(Can the hypothesis on α be deleted?)

5371. *Proposed by W. A. McWorter, University of British Columbia*

Let R be a ring finitely generated by $x_1, \cdots, x_n$. Suppose that every x_i is nilpotent in R and R satisfies the identity $xyx+yxy=0$ for every x and y which are products of the x_i (x and y monomials in the x_i). Prove that R is nilpotent. (R nilpotent means that for some m every product of m elements vanishes.)

5372. *Proposed by L. Carlitz, Duke University*

Let p be a prime >2. Show that the polynomial $x^{p+1}-\frac{1}{2}x^2+2c$ has linear factors (mod p) if and only if c is a quadratic residue (mod p). Moreover, when linear factors occur each has multiplicity two.

5373. *Proposed by L. Carlitz, Duke University*

Let p be a prime >2.

(a) Show that $x^{p+1}+ax+b$, where $a \not\equiv 0 \pmod{p}$, has no irreducible quadratic factor $\pmod{p}$.

(b) Show that $x^{p+1}+ax^2+b$, where $ab \not\equiv 0 \pmod{p}$, has an irreducible quadratic factor $\pmod{p}$ if and only if $(1-a)b$ is a quadratic nonresidue $\pmod{p}$.

5374. *Proposed by Benjamin Volk, Far Rockaway, N. Y.*

Under what conditions for $\{a_n\}$, $\{b_n\}$ may we find nontrivial A, B such that $\lim \{Aa_n+Bb_n\}$ exists?

5375. *Proposed by Benjamin Volk, Far Rockaway, N. Y.*

Given a sequence of real numbers $\{a_n\}$ satisfying $0<a_1<1$, $a_{n+1}=a_n(1-a_n^2)$. Clearly $a_n \downarrow 0$. For which α is it true that $a_n=O(n^{-\alpha})$? How finely may we estimate the convergence of $\{a_n\}$?

5376. *Proposed by Jürgen Elstrodt, University of Münster, West Germany*

Denote by $\log_p x$ the p-fold iterated natural logarithm of x ($\log_1 x=\log x$, $\log_{p+1} x=\log(\log_p x)$). Abel has shown that the infinite series

$$\sum_{n=N}^{\infty} \frac{1}{n \log n \cdots \log_{p-1} n (\log_p n)^{\alpha}}$$

(N sufficiently great, p a fixed natural number) converges if $\alpha>1$ and diverges if $\alpha \leqq 1$. Now let p increase with n, define $p(n)$ for $n \geqq 3$ by $1<\log_{p(n)} n \leqq e$, and prove (a) the infinite series

$$\sum_{n=3}^{\infty} \frac{1}{n \log n \cdots \log_{p(n)} n}$$

diverges. Prove also (b) the infinite series

$$\sum_{n=9}^{\infty} \frac{1}{n \log n \cdots \log_{q(n)-1} n (\log_{q(n)} n)^{\alpha}}$$

with $q(n)=[\frac{1}{2}p(n)]$ converges if $\alpha>1$ and diverges if $\alpha \leqq 1$. (For real x we denote by $[x]$ the unique integer satisfying $[x] \leqq x < [x]+1$.)

5377. *Proposed by Erwin Just and Norman Schaumberger, Bronx Community College*

In a ring R, each element x satisfies the equation $x=x^{n+1}$ for some integer n. Prove that $x^n y=yx^n$ for each x and y contained in R.

5378. *Proposed by S. W. Williams, Lehigh University*

Is every locally compact, totally disconnected space a 0-dimensional space? (cf. W. J. Pervin, *Foundations of General Topology*, Academic Press, p. 98, #6.)

5379. *Proposed by S. W. Williams, Lehigh University*

A topological space is called KC if every compact set is closed. The no-point compactification of a topological space X at a point x is the one-point compactification of $X-(x)$ at x. Let X and Y be disjoint spaces such that Y is infinite. Define the following topology T on $Z=X\cup Y$: (i) $G\in T$ if $G\subseteq Y$, (ii) $G\in T$ if $Z-G$ is finite. How does the topology T_x of the no-point compactification of (Z, T) at $x\in X$ compare to the topology T_y of the no-point compactification of (Z, T) at $y\in Y$? Show that (Z, T) is sequentially compact iff it is compact. Is there a minimal condition for X such that (Z, T) is KC but not Hausdorff? That (Z, T) is KC but not compact?

SOLUTIONS OF ADVANCED PROBLEMS

Integer solutions of a pair of Diophantine equations

5184 [1964, 326; 1965, 327]. *Proposed by A. Oppenheim, University of Malaya, Kuala Lumpur*

Find the complete solution in rational integers of

$$x^3+y^3+z^3+t^3=2$$
$$x+y+z+t=2.$$

II. *Solution by the proposer.* The complete solution of the given equations may be derived from the known complete solution of the equation

$$X^2+Y^2+Z^2+2XYZ=1. \tag{1}$$

Use the transformation $2X=x-y-z+t$, $2Y=-x+y-z+t$, $2Z=-x-y+z+t$, $2T=x+y+z+t$, which yields integers X, Y, Z, T if the integers x, y, z, t have even sum. Its inverse is

$$\begin{aligned} 2x &= X-Y-Z+T, \qquad & 2y &= -X+Y-Z+T,\\ 2z &= -X-Y+Z+T, \qquad & 2t &= X+Y+Z+T, \end{aligned} \tag{2}$$

so that

$$8(x^3+y^3+z^3+t^3-2)=4(T^3+3TX^2+3TY^2+3TZ^2+6XYZ-4).$$

Hence, if x, y, z, t satisfy the given equations so that $T=1$, then X, Y, Z satisfy (1).

The complete solution of (1) is given by the following rule (See A. Oppenheim, *On the Diophantine Equation* $x^2+y^2+z^2+2xyz=1$, this MONTHLY, 64 (1957) 101.)

Let p, q, r be any integers with greatest common divisor unity such that one of them is equal to the sum of the other two; let u be any integer $\geqq 0$; then X, Y, Z, given by

$$X=\pm\cosh p\theta, \qquad Y=\pm\cosh q\theta, \qquad Z=\pm\cosh r\theta, \tag{3}$$

where $\cosh\theta = u$ and the ambiguous signs are chosen so that their product is -1. Note that, e.g.,

$$\cosh p\theta = \frac{[u + \sqrt{(u^2 - 1)}]^p + [u - \sqrt{(u^2 - 1)}]^p}{2}$$

and similar expressions for $\cosh q\theta$ and $\cosh r\theta$.

It follows that the complete solution of the given equations is given by (2) wherein $T = 1$ and X, Y, Z have the values given by (3).

System of Circles in a Plane

5267 [1965, 194]. *Proposed by Paul Erdös, Israel Institute of Technology*

Let there be given a system of circles of radius 1 in the plane such that every line meets at least one of them. Prove that, for every k, there is a line that meets k of them.

Solution by D. R. Anderson, University of Wyoming. We assume that there is a number k such that no line passes through k circles, and deduce the following consequence: For any number $m > 0$, the set of circles with centers (r, s) satisfying $-m \leqq s/r \leqq m$ is infinite. Suppose this last assertion is false. Then there is a number N such that all these circles lie below the line $y = 3N$. For each integer $j \geqq 1$, there is at least one circle, say C_j, which intersects the line $y = 3(N+j) - 1$. Let (r, s) be the center of C_j. Let θ_j be the angle formed by the two lines through the origin which are tangent to C_j. Then

$$\theta_j = 2 \sin^{-1}(r^2 + s^2)^{-1/2} > 2(r^2 + s^2)^{-1/2}.$$

Now $|s/r| > m$ and $s \leqq 3(N+j)$ imply that $\theta_j > 2m/3(N+j)\sqrt{(m^2+1)}$. Since no circle intersects more than one of the lines of the form $y = 3(N+j) - 1$, the sum of the angles which are formed by each pair of lines which are tangent to a circle and pass through the origin must be at least

$$\frac{2m}{3\sqrt{(m^2 + 1)}} \sum_{j \geq 1} \frac{1}{N + j},$$

which diverges.

This cannot be; for let us denote the angles formed between the tangent lines to C_j and the x-axis by α_j and β_j, with $0 \leqq \alpha_j < \beta_j < \pi$. Then for some N, $\sum_{j=1}^{N} (\beta_j - \alpha_j) > k\pi$. Let $\{\gamma_j\}_{j \leq M}$ be a listing of the angles in $\{\alpha_j, \beta_j\}_{j \leq N}$ such that $\gamma_1 < \gamma_2 < \cdots < \gamma_M$. Then $\sum_{j=1}^{M-1} (\gamma_{j+1} - \gamma_j) \leqq \pi$. Define the multiplicity m_j of the pair (γ_j, γ_{j+1}) to be the number of the pairs (α_i, β_i) such that $\alpha_i \leqq \gamma_j < \gamma_{j+1} \leqq \beta_i$. Since no line passes through more than k circles it is clear that $m_j \leqq k$. Also it is clear that $\sum_{j=1}^{N} (\beta_j - \alpha_j) = \sum_{j=1}^{M-1} m_j(\gamma_{j+1} - \gamma_j)$. Thus

$$\sum_{j=1}^{N} (\beta_j - \gamma_j) \leqq \sum_{j=1}^{M-1} k(\gamma_{j+1} - \gamma_j) \leqq k\pi,$$

which is a contradiction.

By rotating the axes we obtain the following extension: if $m<m'$, the set of circles with centers (r, s) satisfying $m \leqq s/r \leqq m'$ is infinite.

We now prove: if $1 \leqq j \leqq k$, then there are circles $C_1, \cdots, C_j$, and numbers $m_j < m_j'$ such that if $m_j \leqq m \leqq m_j'$ then the line $y=mx$ passes through $C_1, \cdots, C_j$. This is clearly possible when $j=1$. Suppose the assertion has been proved for some value of j. Only a finite number of circles intersect the lines $y=m_jx$ or $y=m_j'x$. Thus the result of the previous argument guarantees that there is a circle with center (r, s) satisfying $m_j \leqq s/r \leqq m_j'$ which is distinct from $C_1, \cdots, C_j$ and does not intersect $y=m_jx$ or $y=m_j'x$. Let C_{j+1} be this circle. Let $m_{j+1} < m_{j+1}'$ be such that $y=m_{j+1}x$ and $y=m_{j+1}'x$ are tangent to C_{j+1}. Then $C_1, \cdots, C_{j+1}$, m_{j+1} and m_{j+1}' satisfy the assertion. When j is taken to be k, we obtain a contradiction to the original assumption, thereby completing the proof.

The proposer remarks that the result is true for any set of circles—the restriction on the radius is not necessary.

Restricting Digits in a Decimal Expansion

5270 [1965, 322]. *Proposed by Michael Fried, Bell Aerosystems, Wheatfield, N. Y.*

Consider decimal expansions of the form $x=.x_0x_1x_2 \cdots$, where $0 \leqq x \leqq 9$, and suppose that there exists some ordered pair of positive integers (j, k) such that $x_{j+lk}=x_{j+(l+1)k}$, $l=0, 1, 2, \cdots$. Call the set of such expressions I. Is $S=[0, 1]-I$ countable?

I. *Solution by Robert Bowen, University of California at Berkeley.* The set $E_m(j, k, n)$ of real numbers in $[0, 1]$ having an expansion $.x_0x_1x_2 \cdots$, with $x_{j+lk}=n$ for $0 \leqq l \leqq m-1$ consists of a finite number of intervals with total length 10^{-m}. Thus $F(j, k, n)=\bigcap_{m=1}^{\infty} E_m(j, k, n)$ has measure zero. As $I=\bigcup_{j,k,n} F(j, k, n)$, a countable union, I has measure zero and S cannot be countable.

II. *Solution by Robin Sibson, King's College, Sutton, Surrey, England.* With j, k positive integers and $0 \leqq n \leqq 9$, let $I(j, k, n)$ be the set of all $x=.x_0x_1x_2 \cdots$ such that $n=x_j=x_{j+k}=x_{j+2k}= \cdots =x_{j+mk}= \cdots$. Then we shall show that $I(j, k, n)$ is nowhere dense in $[0, 1]$. Suppose to the contrary that the interval (y, z) is contained in the closure of $I(j, k, n)$. Choose distinct elements p, q in $I(j, k, n)$, $y<p<q<z$, and suppose p, q agree in their expansions up to and including the tth decimal place. They also agree in all $j+lk$ decimal places, $l=0, 1, 2, \cdots$. Choose a number r, $p<r<q$, whose decimal expansion agrees with those of p and q in at least the first t decimal places; for some positive number l_0, the $(j+l_0k)$-th place in the expansion for r will be arbitrary since $p<q$. Let the $(j+l_0k)$-th place in r be set at $(n-2)$ or $(n+2)$ and designate the remaining places arbitrarily. Then r is not in $I(j, k, n)$, nor can r be a limit point of $I(j, k, n)$, i.e. $I(j, k, n)$ is not dense in (y, z). Thus $I(j, k, n)$ is nowhere dense.

$I=\bigcup_{j,k,n} I(j, k, n)$ is a countable union of nowhere dense sets—and hence of the first category. Its complement S is of the second category and uncountable.

Also solved by J. R. Blum, E. O. Buchman, Louis Comtet (France), Robert Connelly, Roy O. Davies (England), J. E. Doner, N. J. Fine, G. J. Foschini, A. S. Fraenkel (Israel), D. A. Hejhal, R. A. Jacobson, J. C. Mettauer, Norman Miller, J. E. Motzkin, George Piranian, G. F. Schumm, T. I. Seidman, Thomas Zaslavsky, and the proposer.

Direct Product of Groupoids

5271 [1965, 322]. *Proposed by A. J. Goldman, National Bureau of Standards*

A groupoid is a nonempty set closed under a binary operation written multiplicatively. The direct product $G=G_1\times G_2$ of two groupoids consists of the Cartesian product of their sets together with component-wise multiplication. Suppose that G admits such a factorization in which G_1 has the property that, if $x, y\in G_1$, then $xy=y$, while G_2 has a two-sided identity. Let $L(G)$ consist of the left identities of G. Prove that $G_1\cong L(G)$ and that $G_2\cong Ge$ for some $e\in L(G)$.

Solution by Gomer Thomas, University of Illinois. Let 1 be the (unique) two-sided identity of G_2. Clearly $(G_1, 1)\subseteq L(G)$. On the other hand, if $(x_1, x_2)\in L(G)$, then $x_2=x_2\cdot 1=1$, so $L(G)=(G, 1)\cong G_1$.

Let $e=(x_1, 1)$ be any element of $L(G)$. Then $Ge=(G_1x_1, G_2\cdot 1)=(x_1, G_2)\cong G_2$.

Also solved by R. A. Avelsgaard, Robert Bowen, Joel Brawley, Jr., F. S. Brenneman, C. C. Clever, G. D. Crown, R. O. Davies (England), G. G. Ford & E. W. Ewing, D. P. Geller, D. A. Hejhal, G. A. Heuer, James Joseph, H. E. Lahmann (Germany), E. S. Langford, C. C. Lindner, A. E. Livingston, R. E. Maas, I. J. Martinez, J. L. Matucha, R. B. Merkel, J. C. Mettauer, H. C. Mullikin, Gautam Pandya & Jagdish Prasad, Harsh Pittie, J. F. Porter, K. W. Reed, Jr., G. A. Robinson, Azriel Rosenfeld, R. N. Schneider, Robin Sibson (England), M. F. Smiley, W. R. Smythe, Jr., J. F. Standish, Raymond Stockton, Necdet Üçoluk, Albert Wehrly, Thomas Zaslavsky, D. Ž. Djoković (Yugoslavia), and the proposer.

Editorial Note. Merkel and the proposer call attention to the paper by Tamura, Merkel and Latimer, *The direct product of right singular semigroups and certain groupoids*, Proc. Amer. Math. Soc., 14 (1963), pp. 118–123, (especially Theorem 4) for further information about the situation of the problem. (It is noted that Theorem 6 of the paper is incorrect and has been replaced by: A groupoid S is an M-groupoid if and only if S is a right zero band of groupoids S_μ with identity e_μ, $S_\mu=\bigcup_{\mu\in R}S_\mu$, $S_\mu S_\sigma\subseteq S_\sigma$ such that $e_\mu e_\sigma=e_\sigma$.)

Hausdorff Semi-door Spaces

5272 [1965, 323]. *Proposed by S. M. Robinson, Union College*

A topological space X is called a door space if for each subset $A\subseteq X$, either A or $X-A$ is open. We shall call a space X a semi-door space if for each subset $A\subseteq X$, there is an open set O with the property that either $O\subseteq A\subseteq \mathrm{cl}[O]$ or $O\subseteq X-A\subseteq \mathrm{cl}[O]$. Find an example of a semi-door space which is not a door space. Also, prove or disprove the following statement: If X is both a Hausdorff space and a semi-door space, then X is a door space.

Solution by James Joseph, Howard University. I. Let $X=\{0, 1, 2\}$ and let the open sets be $\varnothing$, $\{0\}$, $\{0, 1\}$, X, providing the topology T on X. (X, T) is not a door space, for $\{1\}$ is neither open nor closed, and it is easy to check that (X, T) is a semi-door space.

II. *A Hausdorff semi-door space is a door space.* Suppose x and y are distinct points of X which are not open and let U and V be disjoint open sets about x and y respectively. Consider $A=\{x\}\cup(V-\{y\})$. If O is an open set such that $A\supseteq O$, then $O\cap U=\emptyset$, since $\{x\}$ is not open, and it follows that x does not belong to $\mathrm{cl}[O]$; so $A\not\subseteq \mathrm{cl}[O]$. If O is an open set such that $X-A\supseteq O$, then $O\cap V=\emptyset$, since $\{y\}$ is not open, and therefore y does not belong to $\mathrm{cl}[O]$; so $X-A\not\subseteq \mathrm{cl}[O]$. Therefore, since X is a semi-door space, there is at most one point which is not open. If there are no such points, X is discrete and door; if x is such a point, $A\subseteq X$ is closed if x is in A, and open if x is in $X-A$, which completes the proof.

Also solved by D. R. Anderson, S. Baron, Ralph Bennett, R. O. Davies (England), G. W. Day, S. W. Dharmadhikari (India), G. A. Faschini, Bill Johnson, M. D. Mavinkurve (India), P. S. Schnare, G. F. Schumm, Gomer Thomas, J. Pelham Thomas, Thomas Zaslavsky, and the proposer.

Group Characterizations

5274 [1965, 323]. *Proposed by J. J. Malone, Jr., University of Houston*

Let $(G, +)$ be a finite group of order g such that if $(N, +)$ is a proper normal subgroup of $(G, +)$ of order n there exists no subgroup $(K, +)$ of $(G, +)$ of order k such that $K\cap N=\{0\}$ and $kn=g$. An example of such a group is the cyclic group of order 4. What is the characterization of such groups? In particular, if $(G, +)$ is not cyclic, is $(G, +)$ simple?

Solution by Azriel Rosenfeld, University of Maryland. If G has the property of the problem, no proper normal subgroup of G can have a direct supplement, so that G is directly indecomposable. If the property fails to hold, by the Second Isomorphism Theorem,

$$(K+N)/N \cong K/(K\cap N) = K/\{0\} \cong K,$$

so that $K+N$ has order $kn=g$ and must be all of G; since $K\cap N=\{0\}$, the decomposition $G=K+N$ is unique. If G is abelian, this decomposition is direct; hence for abelian groups the property is equivalent to direct indecomposability, so that by the Fundamental Theorem on Abelian Groups, G (abelian) has the property if and only if it is cyclic of prime power order. If G is nonabelian, however, it may be directly indecomposable and still fail to have the property. For example, the symmetric group S_n is indecomposable for all n since it always has a smallest proper normal subgroup (van der Waerden, *Modern Algebra*, vol. I, pp. 27 and 150); but $N=$ the alternating subgroup and $K=$ any subgroup of the form $\{e, p\}$, p odd (e.g., a transposition) contradict the property. The quaternion group has the property, since it has a smallest proper subgroup, but it is not simple (in fact, it is Hamiltonian—it has proper subgroups and they are all normal).

Also solved by Gomer Thomas.

Mapping an Ultrafilter

5275 [1965, 323]. *Proposed by Hewitt Kenyon, George Washington University*

Using the terminology of Problem 5148 [1964, 1054], suppose that f and g are functions mapping the set X onto the sets Y and Z, respectively; suppose that H is an ultrafilter eventually in X and let F and G be the ultrafilters consisting of maps of f and g, respectively, of subsets of X belonging to H. If F is equivalent to G, is it necessarily true that H is eventually in $\{t|f(t)=g(t)\}$?

Solution by the proposer. No. Let $Y=Z$ the set of integers. Let $X=Y\times Y$, and let f and g be respectively the first and second coordinate projections of X onto Y. Let K be any nontrivial ultrafilter on Y and denote by K_2 the filter on X generated by sets of the form $A\times A$, where $A\in K$. Use the nontrivial nature of K to select such an ultrafilter H on X that $H\supset K_2$ and H is eventually in $\{(m, n)|m\neq n\}$. Clearly the mapped ultrafilters F and G are both equal to K and to each other, but H is not eventually in $X\cap\{(m, n)|m=n\}=\{t|f(t)=g(t)\}$.

Partitioning a Sum of Euler ϕ-functions

5276 [1965, 323]. *Proposed by L. Carlitz, Duke University*

The following formula is a consequence of known results when a is a power of a prime and e is not a divisor of a^t-1 for $0<t<k$:

$$\sum_{cd=k}\mu(c)a^d = \sum_{e|a^k-1}{}' \phi(e) \qquad (k>1).$$

(See L. E. Dickson, *Linear Groups*, p. 20.) Prove that the result holds for all $a>1$, $k>1$.

Solution by Ernest W. Trost, Technikum Winterthur, Switzerland. We prove first the following

LEMMA. *Let N be a natural number and $d_1, d_2, \cdots, d_m$ a set of divisors of N. If $D[N]$ denotes the set of all divisors of N and $(u, v, \cdots)$ is the greatest common divisor of $u, v, \cdots$, then*

$$S[N] = D[N] - \sum_i D[d_i] + \sum_{i<j} D[(d_i, d_j)] - \sum_{i<j<h} D[(d_i, d_j, d_h)] + \cdots$$

is the subset of $D[N]$ whose members do not divide any d_i ($i=1, 2, \cdots, r$).

Proof. Suppose that $t\in D[N]$ divides just r of the d_i. t occurs once in $D[N]$, r times in $\sum D[d_i]$, $\binom{r}{2}$ times in $\sum D[(d_i, d_j)]$, and so on. From

$$\sum_{h=0}^{r}(-1)^h\binom{r}{h} = (1-1)^r = 0$$

it follows that t is not a member of $S[N]$ if $r\geqq 1$.

Next we recall the formulas

$$(a^u - 1, a^v - 1) = a^{(u,v)} - 1, \tag{1}$$

$$\sum_{d|n} \phi(d) = n. \tag{2}$$

Let $p_1^{e_1} p_2^{e_2} \cdots p_s^{e_s}$ be the prime decomposition of k. Putting $N = a^k - 1$ and $d_i = a^{k/p_i} - 1$, we get from (1)

$$(d_i, d_j) = a^{k/p_ip_j} - 1, \qquad (d_i, d_j, d_h) = a^{k/p_ip_jp_h} - 1, \cdots \qquad (i < j < h < \cdots).$$

Applying the Lemma and (2) we have

$$\begin{aligned} \sum_{e|a^k-1}{}' \phi(e) &= a^k - 1 - \sum_i (a^{k/p_i} - 1) + \sum_{i<j} (a^{k/p_ip_j} - 1) - \cdots \\ &= \sum_{d|k} \mu(d) a^{k/d} - (1-1)^k = \sum_{d|k} \mu(d) a^{k/d}, \end{aligned}$$

completing the proof.

Also solved by M. G. Greening (Australia), D. R. Hayes, Wolfgang Schwarz (Germany), and the proposer.

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: K. O. MAY, University of California, Berkeley and E. P. VANCE, Oberlin College

Materials intended for review should be sent directly as follows: Books: R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457. Programmed Materials: K. O. May, Univ. of California, Berkeley, Calif. 94704. Films: E. P. Vance, Oberlin College, Oberlin, Ohio 44074.

On Formally Undecidable Propositions of Principia Mathematica and Related Systems. By Kurt Gödel. Translated by B. Meltzer with Introduction by R. B. Braithwaite, Basic Books, New York, viii+72 pp., $3.00.

Because Gödel's results are of such extreme importance, the following review consists of a summary and outline of the work.

This is a nice translation of Gödel's famous paper "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I." The Introduction by Braithwaite contains an exposition and some interesting commentary on the methods used in the paper.

It is probably correct to say that the most comprehensive formalizations of mathematics are, on the one hand, the system of Principia Mathematica (PM), and on the other hand, the axiomatic set theory of Zermelo-Fraenkel (ZF).

These systems are both so extensive that it would appear extremely plausible that *all* mathematical questions expressible in the formalism of the system can be decided on the basis of the axioms of the system. Gödel's great discovery is that this is false! Indeed, Gödel shows that for both these systems, as well as for a large and significant variety of related systems, there are relatively simple problems in the theory of ordinary whole numbers which are undecidable in the system—more precisely, there is a number-theoretic sentence X such that neither X nor its negation $\sim X$ can be proved in the system. (Actually, for purely technical reasons, Gödel's proof is carried out in a system P consisting of PM with the adjunction of the Peano postulates for arithmetic, but it is obvious how the proof can be modified for PM, or Zermelo-Fraenkel set theory.) The proof is based primarily on a highly ingenious modification of Cantor's famous diagonal argument used to prove the existence of a non-denumerable set. The following is a synopsis of the main ideas involved.

We shall let "T" stand for the set of *theorcms* (provable sentences) of the system and "R" for the *refutable* sentences—i.e. those sentences X whose negation $\sim X$ is provable. To say that the system P is *simply consistent* means that the sets T, R are disjoint—i.e. that no sentence is both provable and refutable. As part of the formalism of the system, we have symbols "x", "y", "z" as *variables* ranging over the natural numbers $0, 1, 2, \cdots, n, \cdots$. By a *class-sign* is meant a formula $F(x)$ with exactly one free variable x, and for every natural number n, by $F(n)$ is meant the result of substituting the name of n for (all free occurrences of) x in $F(x)$. We say that $F(x)$ is provable for n if $F(n)$ is provable, and *refutable* for n if the sentence $F(n)$ is refutable. We also say that $F(x)$ *represents* the class of all n for which $F(n)$ is provable. And a class (of natural numbers) is called *representable* in P if there is at least one class-sign $F(x)$ which represents it. If there should be a class-sign $F(x)$ which is refutable for each of of the numbers $0, 1, 2, \cdots, n, \cdots$, but which is such that the sentence $(\exists x)F(x)$ (read "there exists a number x such that $F(x)$") is *provable*, then the system is called *ω-inconsistent*—otherwise the system is called *ω-consistent*. This condition does *not* necessarily imply ordinary (simple) inconsistency, though simple inconsistency certainly implies ω-inconsistency (for in a simply inconsistent system, *all* sentences are provable). The precise content of Gödel's First Incompleteness Theorem is that if P is ω-consistent, then P contains a sentence X which is undecidable in the system—i.e. which is neither provable nor refutable in the system. (A subsequent refinement of Gödel's construction, due to J. Barkeley Rosser, shows the existence of an undecidable sentence under the weaker hypothesis of simple consistency.)

The proof (with some minor modifications) consists first of enumerating all class-signs in a certain sequence $F_1(x), F_2(x), \cdots, F_n(x), \cdots$. We define D to be the class of all n such that $F_n(n)$ belongs to R. (This crucial set D might be called a *diagonal* set.) Now Gödel shows how to construct a class term $F_a(x)$ which, under the assumption of ω-consistency, actually *represents* this set D. This means that for every n:

(1) $$n \in D \leftrightarrow F_a(n) \in T.$$

By definition of D, (1) can be stated equivalently:

(2) $$F_n(n) \in R \leftrightarrow F_a(n) \in T.$$

Since (2) holds for every n, we may substitute a for n and obtain

(3) $$F_a(a) \in R \leftrightarrow F_a(a) \in T.$$

Thus the sentence $F_a(a)$ belongs either both to T, R or to neither. Assuming P ω-consistent, it is also simply consistent (as remarked previously), so $F_a(a)$ cannot belong both to T and R. Therefore this sentence belongs neither to T nor R—i.e. it is an undecidable sentence of the system.

Actually, Gödel obtains an even more striking incompleteness property, which curiously enough is not as popularly known. Gödel's undecidable sentence G is of the universal form $(\forall x)F(x)$ (read "for every natural number x, $F(x)$"). This sentence $(\forall x)F(x)$ is not provable, and yet every one of the infinitely many sentences $F(0)$, $F(1)$, $\cdots F(n) \cdots$ is provable! That is to say, for each particular n, there exists a proof for $F(n)$, but there exists no single proof for the sentence $(\forall x)F(x)$. This situation is sometimes paraphrased by saying that the system is ω-incomplete.

In our synopsis of Gödel's argument for the First Incompleteness Theorem, we have given no indication of the proof of the key fact that the set D is representable. This proof, though elementary in principle, is extremely involved in actual execution. It is here that the idea of arithmetization of syntax comes into play. Each formal expression of the system is assigned a certain natural number—the so called *Gödel number* (as termed by subsequent authors). Then any so called metamathematical statement—i.e. any statement about the formal expressions of the system—has a counterpart in a sentence *of* the system about the natural numbers (viz., the Gödel numbers of the expressions).

This brings us now to Gödel's Second Incompleteness Theorem. Consider the metamathematical statement that the system is (simply) consistent. This—like every other metamathematical statement—has an arithmetic sentence as its counterpart; this sentence is called "Consis." Can this sentence Consis be proved in the system? The answer is given in Gödel's Second Incompleteness Theorem, which asserts that if the system is consistent, then the sentence Consis is not provable in the system. The proof is based upon the following facts. The metamathematical statement that Gödel's undecidable sentence G is not provable, also has its arithmetical counterpart. It turns out that this counterpart is the very sentence G itself! Thus the metamathematical statement that the consistency of the system *implies* the unprovability of G has as its arithmetic counterpart the sentence Consis$\supset G$. And the proof of this metamathematical fact can all be formalized (no easy task!) to yield a formal proof of the sentence Consis$\supset G$. Hence if Consis were provable, then G would be provable, which by Gödel's First Incompleteness Theorem would imply that the system was *incon-*

sistent. (We should remark that for the unprovability of G only simple consistency was needed by Gödel; ω-consistency was needed to show that G is not *refutable.*) Thus, if the system is consistent then the sentence Consis is not provable in the system.

This theorem has been sometimes misleadingly paraphrased as saying that if P is consistent, then the consistency of P is not demonstrable in P. We say *misleading* for the following reasons: It is a precise and true fact that the particular sentence Consis is not provable in P. And one might say that in a certain intuitive sense, Consis expresses the consistency of P, but this fact is *not* precise. Indeed, could there not be another sentence Consis′ which also, in some intuitive sense, expresses the consistency of P, yet which *is* provable in P? The situation thus calls for a careful analysis of just what it means for an *arithmetic* sentence to express the *metamathematical* notion of consistency. Such questions have been studied by subsequent authors, notably Solomon Feferman.

In conclusion, let us remark that there are some who regard Gödel's Second Incompleteness Theorem as destroying all hopes of fulfilling Hilbert's formalistic program of finding a so-called *finitary* consistency proof for mathematics. It may be of interest that Gödel himself expressly states his belief (towards the very end of the paper) that this is not necessarily the case. His point is that because a consistency proof is finitary, it does not necessarily follow that it has to be formalizable in PM or ZF.

R. M. SMULLYAN, Yeshiva University

Representations of the Rotation and Lorentz Groups and their Applications. By I. M. Gelfand, R. A. Minlos, and Z. Ya. Shapiro. Macmillan, New York, 1963. viii+366 pp. $10.00.

Lie group theory has two marvelous properties: detailed study of the simplest examples is an extremely good guide to the general theory, and there are very close relations to physics. The simplest interesting examples from the mathematical point of view are the rotation and Lorentz groups, illustrating the difference in behavior between compact and non-compact groups. Of course, these are also the most useful groups in the applications of Lie group theory to physics.

This book gathers together in explicit form the facts about the representations of these two groups. It should be useful and accessible to anyone interested in Lie group theory or its applications. Although there are valuable hints as to possible generalizations scattered throughout, the generalities of group theory are purposely avoided. Another special feature is a chapter on relativistically invariant wave equations, which is considerably clearer than the treatment in most books on quantum mechanics.

There seem to be numerous misprints. For example, the Klein-Gordon equation is called the "Clebsch-Gordan" equation.

R. HERMANN, Argonne National Laboratory

Non-Linear Mathematics. By Thomas L. Saaty and Joseph Brain. McGraw-Hill, New York, 1964. 381 pp. $12.50.

This book, based on the teaching experience of the writers, treats a number of subjects currently of great importance in science. While they all share a common disease, non-linearity, the cure is different in each case. In each the treatment is thorough, up-to-date references are provided, and the problems are excellent. It is difficult, however, to estimate the level at which the book can be used; judging from the reviewer's teaching experience, probably not below the graduate school.

The subjects treated are these: linear and non-linear transformations in infinite-dimensional space, solution of algebraic and transcendental equations, optimization, non-linear differential equations, control theory, non-linear prediction theory.

HARRY POLLARD, Purdue University

Mathematical Discovery, On understanding, learning, and teaching problem solving, Volume II. By George Polya. Wiley, New York, 1965. xxii+191 pp. $5.50.

In Chapters 7 through 13 the author continues his detailed analysis of the art of problem solving. Chapter 14 is devoted to some very sensible views on teaching high school mathematics. Chapter 15 presents some research type problems on the classroom level.

Although it is written with high school mathematics in mind, the principles are universal and apply to all levels. A student or teacher at any level can find much that is interesting and valuable in the entire volume.

H. S. ZUCKERMAN, University of Washington

Linear Partial Differential Operators. (Grundlehren der mathematischen Wissenschaften, Band 116.) By Lars Hörmander. Academic Press, New York, 1963. vii+284 pp. $10.50.

This book is divided into three parts. Part 1 contains a chapter summarizing the theory of distributions and a chapter on the special spaces of distributions to be considered. Part 2, on differential operators with constant coefficients, has chapters on the existence of solutions, the interior regularity of solutions and the Cauchy problem. Part 3 considers differential operators with variable coefficients. A brief chapter on differential equations with no solutions is followed by chapters on operators of constant strength, operators with simple characteristics, the Cauchy problem, and a concluding chapter on elliptic boundary value problems. The treatment is thorough with relatively abbreviated proofs. In the preface the author states that his aim is to give a systematic study of questions of existence, uniqueness, and regularity of solutions of partial differential equations and boundary problems. This claim is met.

STEPHEN HOFFMAN, Trinity College

Engineering Systems Analysis. By A. G. J. MacFarlane. Addison-Wesley, Reading, Mass. 1965. 272 pp. $8.50.

This is an undergraduate, or first-year graduate, text on the automatic control of electrical and mechanical devices. It deals mainly with the mathematical techniques used in modelling these devices and the stability of those described by linear and non-linear differential equations. There is a short section on Pontryagin's Maximum Principle. The section on computational methods in controller design is somewhat brief.

J. J. FLORENTIN, Brown University

Mathematical Methods in Reliability Engineering. By Norman H. Roberts, McGraw-Hill, New York, 1964. 300 pp. $12.50.

This text describes the mathematical techniques of reliability analysis for those who want to apply the theory, but also want a deeper understanding of the principles involved. The presentation is by way of examples, and its special virtue is that the author can make the principles clear while conveying effectively the results of his extensive practical experience.

The main sections are on the reliability of switching circuits and redundancy techniques, probabilistic models of systems with interacting components, and testing to find the values of parameters in these models. Testing is the application of estimation and experimental design techniques under the complicated constraints of real problems, and this lengthy section shows the author at his best. There is also a good introduction to sampling and estimation theory, and a brief account of maintenance theory.

J. J. FLORENTIN, Brown University

Lectures on Differential Geometry. By S. Sternberg. Prentice-Hall, Englewood Cliffs, N. J., 1964. xi+390 pp. $12.00.

This book is both fascinating and exasperating. On the one hand it is the first text to give a more or less unified treatment of differential geometry, Lie groups, the calculus of variations and topics in the theory of principal bundles. On the other hand it becomes, in spots, a morass of typographical errors, bad editing and doubtful exposition. There is an extraordinary amount of information packed into this relatively small tome but one must beware of references to lemmas which don't exist (p. 127), redundant paragraphs (p. 249, 250) and exercises which might well be termed theorems.

Chapter One is a good, but very concise, presentation of the main results of tensor and exterior algebra.

Chapter Two introduces differentiable manifolds, their tangent spaces, principal bundles, vector fields and Lie derivatives. Some stiff analysis concerning the density of critical values of differentiable maps (Sard's theorem) and manifold embedding theorems (via partitions of unity) immediately follows the definition of a manifold. The concept of a tangent vector occurs as an equiva-

lence class of curves and is later linked to the concept of a derivation. Tensor fields are treated in terms of principal bundles. The chapter ends with a section on Lie derivatives.

In Chapter Three one finds manifold calculus,—the differential calculus of exterior forms and the integral calculus of chains and densities. The concepts of orientation, cohomology and mapping degree as well as the theorems of Stokes, Poincaré, Frobenius and Darboux are dealt with in a brief, but generally well-organized, fashion. Hamiltonian dynamics is treated in a coordinate-free manner in the final section.

Chapter Four centers on the calculus of variations. The necessary and the sufficient conditions for extremals of Euler, Lagrange, Legendre and Jacobi appear in a greatly altered but intrinsic form. In the Riemannian case geodesic sprays, the exponential map, completeness theorems and isometries are discussed. Much of the exposition of this chapter seems rather forced.

Chapter Five, on Lie groups, covers invariant vector fields, related Lie algebras, one-parameter subgroups, closed subgroups, adjoint representations, invariant metrics and structure equations (in several forms).

The group theory of Chapter Five is applied in Chapter Six to the group of euclidean motions. In particular the vector-valued forms of the preceding section lead to the equations of structure of euclidean space. These in turn give rise to embedding theorems and Riemannian manifolds. Curve theory including Frenet frames, homotopy, mapping degrees and convexity is then discussed. A treatment of the second fundamental form, Gauss' *theorema egregium* and two results concerning the embedding of n-dimensional manifolds in E^{n+1} follow. Finally surfaces in E^3 are considered. The Gauss-Bonnet theorem, the Euler-Poincaré characteristic and Betti numbers are mentioned as well as bounds on the lengths of geodesics.

The last chapter takes up the problem of general principal bundles, their reductions, connections on them and their relation to holonomy groups. By specialization this leads to the concept of a G-structure which is a special submanifold of the frame-bundle of a manifold. This concept includes almost every structure over a manifold of geometric interest,—complete parallelism, orientation, conformal, almost complex, etc. Many of the basic theorems of the preceding work now appear as special solutions to the problem of deciding when two G-structures are locally equivalent. The rest of the chapter discusses this general problem using the techniques of reduction and prolongation and the device of the structure function (analogous to the structure-constants of a Lie group). The study of connections on G-structures yields torsion and curvature forms.

In view of the tremendous scope of this book, the importance of its subjects and the dearth of comparable texts one must commend the effort highly, but I certainly hope that the prospective reader will take the self-analysis in the author's preface seriously and that more care will be taken in editing future editions.

J. R. VANSTONE, University of Toronto

Foundations of General Topology. By William J. Pervin. Academic Press, New York, 1964. xi+209 pp. $7.95.

This is a suitable textbook for a senior or first year graduate level course in point-set topology in which the emphasis is analytic rather than geometric. We shall have some comments on the advisability of this at the end of the review. The book begins with two chapters (35 pages) on set theory and cardinal and ordinal numbers. In order to get started, it is assumed that the student is acquainted with various unstated "logical laws". Once over this hurdle, the exposition proceeds smoothly and most students will probably be able to ignore the closing admonition that, ". . . the intuitive set theory we have used is actually not correct."

The next three chapters (3, 4, and 5) discuss Topological Spaces, Connectedness, Compactness, Continuity, Separation and Countability. The treatment is standard and more or less in the spirit of Kelley. The relation of these concepts to familiar properties of the real line is not handled too well. A reference to Dieudonné is surely not the proper way to dispose of the fact that the connected subsets of the line are precisely the intervals. However, it is pointed out that this implies the intermediate value theorem. Also one must look very carefully to piece together the Bolzano-Weierstrass Theorem (p. 58, p. 71, and p. 81). Finally, the reader is left to his own devices with regard to the Heine-Borel Theorem. Although it is proved that a continuous function preserves compact sets, it is nowhere mentioned that this implies the maximum value theorem of calculus.

The next two chapters (6 and 7) discuss metric and complete metric spaces. There is a rather nice but brief discussion of Hilbert spaces, Fréchet spaces, and the space of continuous functions on the unit interval with the sup metric. Surely, though, the statements of 6.3.3 and 6.5.1 miss the point—the point being that in this context local compactness is equivalent to finite dimensionality. The completion of a metric space is constructed and the Baire category theorem is proved.

The next two chapters (8 and 9) are on Product Spaces, Function Spaces and Quotient Spaces. Again the treatment is standard. Ultrafilters are introduced to prove the Tichonov Theorem. (Nets are never mentioned except in side remarks.) In function spaces, the topology of pointwise convergence, the compact-open topology, and the topology of uniform convergence on compact sets are considered. It is shown that the second and third coincide on bounded continuous functions if the range space is metric, but one never finds out when the continuous functions form a closed subset. Equicontinuity is not mentioned, so one does not get any form of Ascoli's Theorem.

Chapter 10 discusses paracompactness and metrization theorems (Urysohn and Nagata-Smirnov) while Chapter 11 is on uniform spaces. In both of these chapters the treatment seems well suited to the intended level of the book. For the existence of completions of uniform spaces, the reader is referred to Bourbaki. It is not mentioned that topological groups have natural uniform structures which make continuous homomorphisms uniformly continuous.

Why, then, is this generally good and competent book so unsatisfactory? I think the answer lies in the treatment of topology as an abstraction and formalization of certain procedures that occur in analysis, and even this aspect of the subject is badly motivated as the above comments indicate. One could, for example, take the position that the point of topologies is to tell one when sequences converge. Function spaces then appear interestingly as important examples where sequences are not sufficient and one must go to nets to characterize the topology. Since the characterization of topologies by sequences or nets is not considered, this cannot be done here.

From another point of view, all of these comments are irrelevant, because what is lacking—and not only in this book—is the *spirit* of topology, either as a classical or a modern subject. Classically, topology is a branch of geometry, but there are no pictures in this book. The circle does not even appear, much less the fact that the torus is either the product of two circles, or the quotient of a square with sides identified properly, or a suitable subvariety of four-dimensional space. The word "homeomorphism" is barely mentioned and the word "homotopy" exactly once in a side remark with no definition.

In the modern spirit, one would view topology as an example of category theory (i.e., the point of topologies is to enable one to recognize continuous maps), but there are no diagrams in this book. The universal mapping properties of the product topology and the product uniformity are mentioned in exercises, but not those of quotient spaces, of the Stone-Cech compactification, or of completions; and, the adjointness relation between products and function spaces is never even hinted at. All of these things are, of course, the central facts of life in the category of topological spaces.

All-in-all, I would say that if this book represents what analysts want in point-set topology courses, then we must see to it that such courses are taught only by topologists.

JOHN W. GRAY, University of Illinois

Differential and Riemannian Geometry. By Detlef Laugwitz. Translated by Fritz Steinhardt. Academic Press, New York, 1965. xii+238 pp. $8.50.

This is a good introduction to the essential ideas and techniques of differential geometry. Although it is relatively short, the book has a tremendous scope. The reader comes in contact with ideas of interest in contemporary research and with topics not ordinarily found in an introductory work.

In Chapters I and II, taking up about one-third of the text, the author discusses classical differential geometry of curves and surfaces in three-dimensional Euclidean space. Tensor notation is used throughout, with the notion of a tensor defined somewhat rigorously as a multilinear form in Chapter III. There is no discussion of the calculus of differential forms, because of "lack of space."

Chapter III is devoted to a development of tensor calculus, to the geometry of a space with affine connection and to the foundations of Riemannian geometry. Of special interest here is the discussion of the holonomy group of a connection, including the result that, with a suitable choice of coordinates, the

holonomy group at any point of a Riemannian space is an orthogonal group. The author also discusses the projective equivalence of two affine connections, and the notion of completeness of a manifold. He does not give a proof of the equivalence of conditions for completeness.

Chapter IV begins with a discussion of spaces of constant curvature, including a proof of the existence of complete Riemannian manifolds of constant negative curvature. Of special interest is the section on applications of Riemannian geometry to analytical dynamics. This is followed by the definition of a Finsler space, a discussion of the indicatrix of the Minkowski metric, and the characterizations of Riemannian spaces due to Helmholtz and Weyl. The chapter closes with some results on Finsler spaces and nonlinear connections.

Chapter V is a brief one, containing selections from differential geometry in the large. Included here are a proof of Fenchel's Theorem on the total curvature of closed space curves, characterizations of the sphere, the Gauss-Bonnet Theorem, and the proof due to Herglotz, using integral formulas, of the theorem of Cohn-Vossen on the congruence of isometric ovaloids. The book closes with three short appendices on the history of differential geometry, existence and integrability theorems for differential equations and a summary of formulas.

The book is quite readable. There are adequate problem sections, with hints or references given for the more difficult exercises. There seem to be few errors or misprints. The reviewer has used the book with several senior physics students on an independent study basis and can affirm that the material presented is well within the grasp of such students. Because of its unusual features in Chapters III and IV, its brevity and its organization, the book represents a useful addition to the existing literature in elementary differential geometry.

G. F. Feeman, Williams College

Mathematics and Logic in History and Contemporary Thought. By Ettore Carruccio. Translated from the Italian by Isabel Quigly. Aldine, Chicago, 1965. 398 pp. $8.75.

This is not a history text nor even a concise history for the layman. Rather it is an historically ordered collection of chapters on subjects chosen "to show the development of the most important fundamental concepts of mathematics, and particularly the contribution made by mathematical thought to the evolution of logic . . . in order to emphasize subjects which seem fundamental in the training of those who will do research in the history and philosophy of mathematics." References (even for quotation of non-Italian authors) are almost entirely to Italian secondary sources, and the student who does not read Italian will find the book a useful survey of the work of Carruccio (twenty-six articles by him from 1927 to 1957 are listed) and of the Italian authors he prefers (F. Enriques is his favorite, and G. Loria is almost ignored). The bibliography has not been revised for English readers, so that for them the footnotes are not helpful. Such obvious authors as Neugebauer, van der Waerden, Bell, Hofmann, and Struik are not cited. Chinese mathematics is discussed without taking ac-

count of Needham's work. The translation is smooth, though a few mathematical terms are mistranslated. The original work, published in 1958, is undoubtedly useful to Italian students of philosophy, but English readers would have found more valuable a selection of the author's original papers on history and foundations.

KENNETH O. MAY, University of California Berkeley

Elementary Differential Topology. By James R. Munkres. Lectures given at M.I.T. fall 1961. Annals of Mathematics Studies No. 54. Princeton Univ. Press, Princeton, 1963. xi+107 pp. $3.00.

One of the prime techniques in differential topology is that of approximating a map by a smoother one with similar properties. The known facts concerning the possibility of such approximations, though often used, are rarely proved. The eminently readable book under review presents a clear, unhurried exposition of these facts, in a highly successful attempt to make them more accessible to the student of differential topology who will need them.

The first chapter, entitled "Differentiable manifolds," has as its goal the twin theorems that each C^1 differentiable structure on a manifold contains an infinitely differentiable structure, and that each once differentiable immersion, embedding, or diffeomorphism can be approximated by an infinitely differentiable one. The second chapter, entitled "Triangulations of differentiable manifolds," deals with the existence of smooth triangulations, the existence of (piecewise) linear approximations to smooth immersions or embeddings defined on triangulated manifolds, and the consequent possibility of using small alterations to fit together two embedded complexes so that their intersection is a subcomplex.

Exercises and problems, of varying difficulty, abound. The diligent student will find them a valuable aid to the understanding. It is to be hoped that more such informal tracts on advanced material will see fit to do the reader the same favor of allowing him to try his own hand at the game.

F. E. J. LINTON, Wesleyan University and the University of Chicago

Elements of Complex Variables. By Louis L. Pennisi. Holt, Rinehart, and Winston, New York, 1963. x+459 pp. $7.95.

This undergraduate text is carefully written—in places too carefully written. It shoulders the burden of being everything from a student's introduction to rigor to an engineer's text. Even at that it succeeds remarkably well. Chapter headings are: 1. Complex Numbers; 2. Point Sets, Sequences, Mappings; 3. Analytic Functions; 4. Elementary Functions; 5. Integration; 6. Power Series; 7. Calculus of Residues; 8. Conformal Mapping; 9. Theory of Flows.

Consideration of analytic functions begins at a snail's pace by proving all the theorems about limits and continuity the student has presumably already seen in calculus, with the same proofs. Once the derivative is defined, the subject

begins to flow smoothly. The chapter on elementary functions includes the standard multiple-valued functions, done very patiently and very well indeed. In the integration chapter, the author spells out the geometric preliminaries with precision, but when an integral appears he temporarily retreats. His way of asserting that various forms of the integral are the same number seems to call for a theorem that appears nowhere in the book. Rigor returns with the Cauchy integral theorem, the proof following Ahlfors.

With the difficulties over, the chapters on power series and calculus of residues are designed to help the reader develop his own powers of calculation. The author includes a wealth of examples. The last two chapters seem standard, except that the author includes enough about the Joukowski airfoil to make it intelligible.

The great number of exercises is one of the virtues of the book. The author's informal remarks are excellent—some observational, some cautionary. The virtue of patience in exposition is also the book's only major fault. There is much space devoted to preliminary details that could be fitted into the exercises. This may account for the book's uncomfortable size: too long for a semester, not enough content for a year.

STEPHEN PUCKETTE, University of the South

Partial Differential Equations. By Günter Hellwig. Blaisdell, New York, 1964. 263 pp. $10.00.

This book is an excellent translation; the mathematical English is very good, and the text as a whole reads smoothly. The original work appeared in German in 1960. The book covers a surprisingly wide variety of material, some classical and some modern. Although it omits some topics one might expect in such a work—for example, the method of separation of variables—the penetrating and rigorous treatment of the material selected gives the reader more than a passing acquaintance with it. Some recommending features: 1) Enough detail is provided in somewhat involved arguments so that the reader who sees them for the first time will not be at a complete loss. 2) Choice of examples and counterexamples and their treatment to illustrate the theory is excellent. 3) The author cites original sources when practical and otherwise reports on surveys which review the literature in specific areas. 4) There are few typographical errors. Probably the worst of these are three at the top of p. 51, and they are not serious. 5) A few (but not enough) problems are scattered throughout the text with solutions of the more difficult ones appearing at the end of the book. One distracting feature: the author's method of referencing results proved earlier in the book is cumbersome since the part, chapter and section numbers are not printed at the top of each page. Prerequisites include a good knowledge of advanced calculus, and the theory of ordinary differential equations, some function theory and a little functional analysis. This is a highly valuable reference work for the serious student in this branch of analysis.

R. L. BORRELLI, Harvey Mudd College

Abstract Harmonic Analysis. By Edwin Hewitt and Kenneth A. Ross. Volume I: Structure of topological groups; integration theory; group representations. Springer-Verlag, Berlin, 1963. viii+519 pp. $19.00.

The book under review is the first half of a projected two volume monograph on abstract harmonic analysis. The second volume (we are told in the preface) "will treat harmonic analysis on compact groups and locally compact Abelian groups, in considerable detail." The present volume "gives all of the structure of topological groups needed for harmonic analysis as it is known to us; it treats integration on locally compact groups in detail; it contains an introduction to the theory of group representations." Three appendices recall relevant facts about Abelian groups, topological linear spaces, and normed algebras.

In a sense, this volume lays out the lubricants required for the smooth operation of the machinery to be found in the second. One could equally well maintain that this volume lays the foundation on which the edifice of the second is to rest. In either case, it seems indisputable that the absence of much serious representation theory in the present volume and the presence of so much material of prerequisite character will disappoint a fair number of readers. Equally indisputable, however, and much to the book's credit, is the comprehensive treatment of the subjects that are covered.

There are many exercises in sections entitled *Miscellaneous theorems and examples*, and copious historical notes following most paragraphs. The presence of an index of symbols is a thoughtful service.

F. E. J. LINTON, Wesleyan University and the University of Chicago

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor Norman Haaser, University of Notre Dame, represented the Association at ground breaking ceremonies of a new science building, which will house the Department of Mathematics at Taylor University, on November 29, 1965.

Professor Rufus Oldenburger, Purdue University, was elected a Fellow of the American Society of Mechanical Engineers.

Brigham Young University: Dr. R. V. Skarda, California Institute of Technology, and Mr. L. K. Tolman, University of New Mexico, have been appointed Assistant Professors.

University of British Columbia: Mr. R. S. Booth and Mr. S. D. Promislow have been appointed Lecturers; Associate Professor C. A. Swanson has been promoted to Professor; Assistant Professors C. W. Clark and Roy Westwick have been promoted to Associate Professors.

San Fernando Valley State College: Dr. D. H. Potts, University of California at Berkeley, has been appointed Professor; Mr. G. R. Grainger, Planning Research Corporation, Los Angeles, California, and Dr. Stoddart Smith, Jr., Douglas Aircraft Company, Santa Monica, California, have been appointed Assistant Professors; Associate Professor J. W. Blattner has been promoted to Professor.

U. S. Air Force Academy: Major L. G. Campbell, formerly Head of Mathematics Department at USAFA Prep School, and Major R. C. Rounding have been appointed Associate Professors; Associate Professor J. B. MacWherter has been promoted to Professor and appointed Head of the Mathematics Department; Assistant Professor R. R. Erbschloe has been promoted to Associate Professor; Major C. A. Wurster has been promoted to Assistant Professor.

Winthrop College: Professor B. G. Hodges, Illinois State University, has been appointed Associate Professor; Mr. H. H. Crockett, Duke University, has been appointed Assistant Professor.

Yeshiva University: Dr. Joseph Lewittes, Harvard University, has been appointed Assistant Professor; Dr. C. R. Patt has been promoted to Assistant Professor.

Dr. Richard Bellman, Rand Corporation, Santa Monica, California, has been appointed Professor of Mathematics, Medicine, and Engineering at the University of Southern California.

Dr. B. A. Chartres, Brown University, has been appointed Associate Professor of Computer Science and Associate Director of the Computer Science Center at the University of Virginia.

Professor Eugene Lukacs, Director of the Statistical Laboratory of the Catholic University, is on leave during the academic year 1965–66. He is spending part of the time at the University of Vienna and part at the Sorbonne.

Mr. Martin Peres, New York University, has been appointed Assistant Professor at Mercy College.

Professor R. F. Rinehart, Case Institute of Technology, has been appointed Senior Professor at the U. S. Naval Postgraduate School.

Sister Mary Colum O'Donnell, Rosary College, has been appointed Chairman of the Physics Department.

Professor C. R. Adams, Brown University, died on October 15, 1965. He was a member of the Association for 46 years.

Dr. Charles E. Clark, System Development Corporation, Santa Monica, California, died on June 16, 1965. He was a member of the Association for 16 years.

Mrs. Marjorie L. French, Topeka Public Schools, Kansas, died on July 5, 1965. She was a member of the Association for 6 years.

Miss Bertha I. Hart, Birmingham, Alabama, died on September 7, 1965. She was a member of the Association for 42 years.

Professor Emeritus W. R. Longley, Yale University, died on February 23, 1965. He was a charter member of the Association.

Mrs. Alice B. Rabon, Brookland-Cayce High School, Cayce, South Carolina, died on June 23, 1965. She was a member of the Association for 7 years.

Sister Mary Job Ternes, Rosary College, died on February 1, 1965. She was a member of the Association for 5 years.

Rev. G. W. Walker, Franklinville, New York, died on November 17, 1965. He was a member of the Association for 19 years.

Associate Professor Emeritus S. D. Zeldin, Massachusetts Institute of Technology, died on November 2, 1965. He was a member of the Association for 48 years.

HOW ABOUT A CAREER WITH MATHEMATICS

An eight-page leaflet entitled "How about a career with mathematics" is published annually by the Committee on High School Contests for distribution to all participants in the contest. The contest is supported by MAA, the Society of Actuaries, and Mu Alpha Theta.

This year the Committee has had additional copies of this leaflet printed, to be supplied free of charge to any MAA members who might wish to have them for distribution to students. Anyone wishing copies of the leaflet "How about a career with mathematics" should write to: MAA, SUNY at Buffalo (University of Buffalo), Buffalo, N. Y. 14214.

REQUEST FOR INFORMATION FROM CEM

We have received the following letter from Dr. A. N. Feldzamen, Executive Producer, Committee on Educational Media:

On the morning of Monday, December 29, 1952, John von Neumann gave a lecture on Game Theory at a meeting in St. Louis of the American Association for the Advancement of Science. The Mathematical Association of America was also meeting there and his talk had auditors from both meetings. Someone in the audience—identity unknown—tape-recorded this talk. Can you help us locate this tape?

The Committee on Educational Media of the Mathematical Association of America is preparing a film on the life and achievements of John von Neumann. We are trying to trace down this recording and other recordings, and gather pictures and other information about him. Photographs are especially wanted; all such will be carefully handled and returned to the sender. If you can help us, please write or call (collect) Miss Patricia Powell, 344 West 12th Street, New York, N. Y. 10014—(212)243-5318, or Dr. A. N. Feldzamen, Committee on Educational Media, P.O. Box 2310, San Francisco, Calif. 94126—(415)362-7582.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

NOVEMBER MEETING OF THE MINNESOTA SECTION

The annual fall meeting of the Minnesota Section of the MAA was held on November 13, 1965 at Moorhead State College, Moorhead, Minnesota. Professor Marion V. Smith, Moorhead State College, presided at the morning session. Section Chairman, Professor Frank Wolf, Carleton College, conducted the afternoon session. There were 94 persons registered for the meeting, of whom 76 were members of the Association.

At the business meeting a report of the MAA meeting at Cornell University, Ithaca, New York, on August 31, 1965, was made by the Section Secretary, Professor W. C. Kalinowski. Professor Seymour Schuster spoke of the availability of certain materials produced by the College Geometry Project at the University of Minnesota. The Minnesota Section congratulated its charter member, Professor Clifford Mills of Sioux Falls, South Dakota. A citation to Professor Paul Rosenbloom was read.

10. *Jacobian matrices in a one dimensional subspace of* $L(E^n, E^n)$, by A. Wayne Roberts, Macalester College, St. Paul, Minnesota.

Let f map an open set U of E^n into E^n. Following Fréchet and the later work of the Nevanlinnas and Dieudonné, we view the derivative of f as a mapping of U into the linear transformations $L(E^n, E^n)$. If a basis in E^n is fixed, a one dimensional subspace of $L(E^n, E^n)$ takes the form $a(p)A$ where $a:U\to Re$ and A is an $n\times n$ matrix. We can show that if rank $A\geq 2$, then $f'(p)=a(p)A$ for every p in U implies that $a(p)=a$ is constant, and hence $f(p)=aAp$.

W. C. KALINOWSKI, *Secretary-Treasurer*

NOVEMBER MEETING OF THE NORTHEASTERN SECTION

The 11-th annual meeting of the Northeastern Section of the MAA was held at Cleveland Hall, Mount Holyoke College, South Hadley, Mass., on Saturday November 27, 1965. In the morning session, Professor Grace Bates of Mount Holyoke College, Chairman of the Section, presided. In the afternoon session, Professor Hartley Rogers of the Massachusetts Institute of Technology, Vice-chairman of the Section, presided. There were 163 persons present including 136 members of the Association.

At the business meeting in the afternoon the following officers were elected for the coming year: Chairman: Hartley Rogers, Massachusetts Institute of Technology; Vice-chairman: Robin Robinson, Dartmouth College; Sec. Treas.: George Best, Phillips Academy.

The following papers were presented:

1. *Current developments in set theory*, by Anil Nerode, Cornell University.

2. *Topological semigroups*, by Haskell Cohen, University of Massachusetts and Louisiana State University.

3. *Applications of probability to analysis*, by Shizuo Kakutani, Yale University.

4. *The Dartmouth time-sharing computing system; Lecture and demonstration*, by J. G. Kemeny, Dartmouth College.

R. S. PIETERS, *Secretary-Treasurer*

NOVEMBER MEETING OF THE PHILADELPHIA SECTION

The fortieth annual meeting of the Philadelphia Section of the MAA was held at West Chester State College, West Chester, Pennsylvania, on November 20, 1965. The Chairman, Professor Russell Remage, Jr., University of Delaware, presided at the meeting. The meeting was attended by 145 persons including 115 members of the Association.

At the business meeting, the following officers were elected: Chairman: Russell Remage, Jr., University of Delaware; Member of the Executive Committee, Martin F. Hubley, Abington School District. Professor R. L. Wilder, University of Michigan, read the names of six charter members of MAA now affiliated with the Philadelphia Section.

The following papers were presented:

1. *The role of the intuition*, by R. L. Wilder, University of Michigan.

2. *Relations which are equivalent with functional equations involving the Riemann zeta functions*, by Fritz Oberhettinger, Oregon State (Visiting Professor, University of Delaware).

3. *Report of Joint Committee on Teacher Certification Standards*, by A. E. Filano, West Chester State; M. F. Hubley, Abington School District; Bernard Jacobson, Franklin and Marshall.

4. *CUPM General Curriculum for Colleges*, by H. O. Pollak, Bell Telephone Labs.

V. V. LATSHAW, *Secretary*

CALENDAR OF FUTURE MEETINGS

Forty-seventh Summer Meeting, Rutgers, The State University, New Brunswick, New Jersey, August 29–31, 1966.

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN, Waynesburg College, Waynesburg, Pennsylvania, April 30, 1966.

ILLINOIS, Saint Dominic College, St. Charles, May 13–14, 1966.

INDIANA, Indiana State University, Terre Haute, May 14, 1966.

IOWA, Central College, Pella, April 15, 1966.

KANSAS

KENTUCKY, University of Kentucky, Lexington, April 29–30, 1966.

LOUISIANA-MISSISSIPPI

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, College of William and Mary, Williamsburg, Virginia, April 30, 1966.

METROPOLITAN NEW YORK, Brooklyn Polytechnic Institute, New York, April 30, 1966.

MICHIGAN, Wayne State University, Detroit, Michigan, April 2, 1966.

MINNESOTA, Macalester College, St. Paul, Minnesota, April 30, 1966.

MISSOURI, University of Missouri at Rolla, Rolla, April 30, 1966

NEBRASKA, Nebraska Center for Continuing Education, Lincoln, April 29–30, 1966.

NEW JERSEY, Ocean Township High School, Oakhurst, May 7, 1966.

NORTHEASTERN

NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.

OHIO, Ohio Wesleyan University, Delaware, April 23, 1966.

OKLAHOMA-ARKANSAS, Oklahoma Baptist University, Shawnee, April 1–2, 1966.

PACIFIC NORTHWEST, University of Victoria, Victoria, British Columbia, June 17, 1966.

PHILADELPHIA, Villanova University, Villanova, November 1966.

ROCKY MOUNTAIN, Colorado State University, Fort Collins, May 13–14, 1966.

SOUTHEASTERN

SOUTHERN CALIFORNIA

SOUTHWESTERN, University of New Mexico, Albuquerque, April 1–2, 1966.

TEXAS, Southern Methodist University, Dallas, April 15–16, 1966.

UPPER NEW YORK STATE, St. Bonaventure University, Olean, May 14, 1966.

WISCONSIN, Wisconsin State University, Eau Claire, May 7, 1966.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, Rutgers, The State University, New Brunswick, New Jersey, August 30–September 2, 1966.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Washington State University, Pullman, June 20–24, 1966.

ASSOCIATION FOR COMPUTING MACHINERY, Ambassador Hotel, Los Angeles, August 30–September 1, 1966.

ASSOCIATION FOR SYMBOLIC LOGIC, Waldorf-Astoria Hotel, New York City, April 4, 1966.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

INSTITUTE OF MATHEMATICAL STATISTICS, Rutgers, The State University, New Brunswick, New Jersey, August 30–September 2, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Americana Hotel, New York City, April 13–16, 1966.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Miramar Hotel, Santa Monica, California, May 18–20, 1966.

PI MU EPSILON

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Iowa, Iowa City, May 11–14, 1966. (Symposium on Numerical Analysis).

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN
MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF
THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 4

PART I

CONTENTS

APRIL 1966

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE TO AUTHORS

More than enough Main Articles, Mathematical Notes, and Classroom Notes have been accepted (or are under consideration) to fill these Sections of the MONTHLY through the December, 1966 issue. Please do not submit any papers for these Sections until after July 1. The June-July issue will contain addresses to which papers may then be submitted.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

CYCLIC ELEMENTS IN TOPOLOGY, A HISTORY

B. L. McALLISTER, South Dakota School of Mines and Technology

By a *Peano continuum* is meant any metric space that is the continuous image of a closed interval. Jordan [**38**] called such spaces continuous curves, but this name lacked universal appeal after 1890, when Peano published his famous example showing that a continuous curve can fill up a plane area [**73**]. By about 1913, Hahn [**27**] and Mazurkiewicz [**57**] had characterized the continuous curve as a locally connected metrizable continuum. There followed a period of great interest and rapid progress in problems relating to the structure of these spaces.

For many purposes it seemed desirable to divide a Peano continuum into "lumps" such that the internal structure of each lump might exhibit a reasonable degree of simplicity, and yet such that the relationships among the lumps might also be relatively simple. The cyclic element theory, initiated in 1926 by G. T. Whyburn, provided such a method of decomposition.

Whyburn, in his dissertation [**101**], had already used a decomposition of a Peano continuum X into three sets: the set of all endpoints of X, the set of all cutpoints of X, and a set N that is the union of the set of all simple closed curves in X. (A point p is called an *endpoint* of X provided that p has arbitrarily small neighborhoods with singleton (i.e. one-point) boundaries. A point p is called a *cutpoint* of (connected set) X iff $X-\{p\}$ is disconnected. A *simple closed curve* is any homeomorph of a circle.) This decomposition was also implicit in work done at about the same time by W. L. Ayres [**3**].

The essential step in the invention of cyclic elements lay in the further subdivision of the set N above into what were called maximal cyclic curves. A Peano continuum C is called *cyclic* provided that for every two points of C there is a simple closed curve in C that contains both points. Then by a *maximal cyclic curve* of a Peano continuum X is meant a set C maximal with respect to being a cyclic Peano continuum contained in X. The maximal cyclic curves of X are also called the *true* (or *nondegenerate*) *cyclic elements* of X, the other (degenerate) cyclic elements being the endpoints of X and the cutpoints of X.

Whyburn first announced his results to the American Mathematical Society in December of 1926, in a brilliant set of three papers. In [**102** and **103**] he characterized cyclicity and maximal cyclic curve in several ways, establishing a degree of simplicity of internal structure for the "lumps." (One of these characterizations, the *cyclic connectivity theorem* is of special importance. It states that a Peano continuum is cyclic if and only if it contains no cutpoint of itself.) In these papers he also proved some basic facts on the relationship of a Peano continuum X to its cyclic elements. For example every simple closed curve in X lies in some cyclic element of X; two cyclic elements of X meet (if at all) in a single point; the number of true cyclic elements of X is countable.

In [**106**] the principal task was to examine the manner in which the cyclic elements of a Peano continuum X combine to form X. Essentially what was shown was that the structure of X relative to its cyclic elements is analogous to

the structure of a *dendrite*, a Peano continuum that contains no simple closed curve. The tools developed to show this analogy are important in their own right, so that we may justify defining them and mentioning the most important theorems here.

In viewing a Peano continuum X as composed of its cyclic elements, the structural analogue of a point is, of course, a cyclic element. The analogue of an arc is a cyclic chain: a subset M of X is called a *cyclic chain* of X between two cyclic elements C_1 and C_2 of X provided that M (1) is connected, (2) is a union of cyclic elements of X, (3) contains C_1 and C_2, and (4) is minimal with respect to properties (1), (2), and (3). Then C_1 and C_2 are called *end elements* of M and the other cyclic elements of X that are contained in M are called *interior elements* of M. The acyclicity of X relative to its cyclic elements is shown, e.g., by the fact [**106**, Theorem 3] that between any two cyclic elements of X there is exactly one cyclic chain.

(A cyclic chain may be viewed either as a collection of cyclic elements or as the union of such a collection. In the late 1920's, the principal tool for dealing with collections of "lumps" of a space was Moore's (still very useful) notion of upper semi-continuous decomposition. But the elements of an upper semi-continuous decomposition had to be pairwise disjoint. Cyclic elements did not fit this scheme very well, because two cyclic elements can intersect; hence Whyburn chose to view the cyclic chains of X as unions of cyclic elements. Several different repairs have been made, so that the structure of X relative to its cyclic elements can actually *be* that of a dendrite, rather than merely presenting an analogy. See, for example, [**28, 43, 64, 145**]. Particularly successful was [**2**]; the "hyperspace" obtained has, however, only very weak separation properties.)

Just as a set M is arcwise connected provided that each two points of M are connected by an arc in M, so also a subset M of Peano space X is called *cyclic chain-wise connected* provided that each two cyclic elements of X in M are connected by a cyclic chain in M. An important class of cyclic chain-wise connected subsets of a Peano continuum X are those connected sets H which are unions of cyclic elements of X. Such sets are now called H-sets; H-sets that are continua are also called A-sets. Whyburn showed that H-sets are cyclic chain-wise connected, arc-wise connected, and locally connected. As a corollary, an A-set of a Peano continuum is a Peano continuum. The closure of any H-set is an A-set.

Whyburn had, in earlier work [**101**], characterized an endpoint of a Peano continuum X as a point p such that p is a cutpoint of no arc in X. He now defined a *node* of X, characterizing it as a cyclic element C of X such that no cyclic chain in X has C as an interior element. Thus *node* is seen to be analogous to *endpoint*. But Whyburn also characterized a node of X as a set which is either a single endpoint or a true cyclic element C such that the complement of the interior of C is connected. This makes a node look like an analogue of a more general non-cutpoint. Of course the reason is that in a dendrite all non-cutpoints

are endpoints, and the "dendritic" relative structure of the cyclic element decomposition behaves analogously.

Another important property of an H-set H of a Peano continuum X is that the intersection of H with any connected subset of X is connected [**108**].

Whyburn's presentation in 1926 had been confined to plane point sets. The topological theorems which made possible the generalization of cyclic element theory to an arbitrary metric space were proved by Ayres [**7**], and Whyburn's paper [**106**] was delayed until 1928 to incorporate the improvement. In the meantime, Ayres had approached the ideas of H-set and node by a different route [**5**]. For each subset K of a Peano space X, he defined a set $X(K)$ called the *arc-curve* of K, namely the union of all the arcs in X that have both endpoints in K. (If K is a single point, $X(K)$ was defined to be K.) If it happens that $X(K)=X$, then K is called a *basic set* of X.

It soon appeared that H-sets and arc-curves were strikingly alike. In fact, Ayres and Whyburn found that the resemblance is perfect: for every nondegenerate K, $X(K)$ is an H-set, and for every H-set of X, there is a set K such that $H=X(K)$. Furthermore, whenever K consists of exactly two points, $X(K)$ is a cyclic chain, and every cyclic chain can be so expressed [**8**, Theorems 1 and 2].

A rather similar relationship exists between Ayres's basic sets and Whyburn's nodes. Letting X be a Peano continuum, but excluding the irrelevant case in which X is cyclic, a subset K of X is a basic set iff for each node C of X, K contains a point of C that is not a cutpoint of X. Iff K contains one such point for each node of X and K contains no other points, then K is an *irreducible* basic set, i.e. a basic set no proper subset of which is a basic set [**8**, Theorems 7 and 9].

Although the theories of arc-curves and basic sets were thus largely absorbed into the cyclic element theory, their distinctive techniques had produced some useful results. In [**4**], Ayres used (modified) arc-curves of a continuous curve X to characterize cyclic continuous curves (and acyclic ones). In [**5**] he showed that arc-curves are locally arc-wise connected; that a component of the complement of an arc-curve H has exactly one limit point in H; that a cutpoint of an arc-curve of X is a cutpoint of X. In [**8**] he proved among other things: that the intersection of two H-sets of X is again an H-set of X; that an A-set A of X is a simple cyclic chain of X if and only if every three points of A lie on some arc of X; that for each three points x, y, and z of X, either there is an arc of X that contains all three points or there is a cyclic element C of X such that $X-C$ contains all three of x, y, z and no two of them lie in the same component of $X-C$.

Ayres's contributions to the extension of cyclic element theory to spaces that are not subsets of the plane have already been noted. In a sense he generalized the spaces under consideration in another way, since he often used the term *continuous curve* to mean a continuous image of an open interval. (Mazurkiewicz [**57**] gives a relevant characterization.)

In describing arc-curves and basic sets as "absorbed" into cyclic element

theory, I should mention that not all of the results of [**8**] on basic sets go over naturally into statements about nodes. For example, every irreducible basic set of X is totally disconnected; the complement of an irreducible basic set is locally connected and consists of a finite number of components [**8**, Theorems 16, 17A, 18].

The cyclic element theory found immediate application to many problems in topology. For example, Whyburn showed in [**104**] that if every true cyclic element of a Peano continuum X is a simple closed curve, then X is a *regular curve* in the sense of Menger, i.e., each point of M is contained in arbitrarily small open sets with finite boundary. Further use of cyclic elements on regular curves was made in [**82**, **107**, and **111**].

A rational curve is slightly more general than a regular one. A Peano continuum X is called *rational* provided each point is contained in arbitrarily small open sets with countable boundary. In [**117**], Whyburn used cyclic elements to show that a Peano continuum in the plane is rational if every subcontinuum of M is a Peano continuum.

Relevant also is Whyburn's contribution [**110**] to the solution of Menger's problem of characterizing a Peano continuum X that is the union of a countable collection of arcs. Whyburn proved that X is a countable arc-sum if and only if all the true cyclic elements of X are countable arc-sums and the endpoints of X are countable. This is probably the first use of a cyclicly extensible and reducible property. (A property is *cyclicly extensible* provided that if it is possessed by each cyclic element of X, then it is possessed by X itself. A property is *cyclicly reducible* when its possession by X implies its possession by each cyclic element of X.)

Whyburn dealt with the more general problem of countable curve-sums for certain types of curves in [**121**]. A little later, Steenrod gave similar results for finite arc-sums [**89**] and finite curve-sums [**88**].

R. L. Moore's *upper semi-continuous decomposition* has already been mentioned. The subject may be considered as part of the theory of closed (continuous) maps (see [**42**] p. 99). We say that a map on a topological space A into a topological space B is *closed* provided that for each closed subset M of A, $f(M)$ is closed. We say that $f\colon A\rightarrow B$ is *monotone* provided that f is continuous and for each point p in B, $f^{-1}(p)$ is a point or a continuum. In [**63**] Moore used cyclic elements to characterize monotone upper semicontinuous decomposition spaces (i.e. spaces that can be images under closed monotone maps) of the 2-sphere as cactoids (i.e. as Peano continua each of whose true cyclic elements is a 2-sphere). Moore also showed that a necessary and sufficient condition that a Peano continuum X in Euclidean 3-dimensional space be the boundary of one of the components of its complement is that X be a cactoid. (For the plane this last result is attributable to Ayres [**10**]. For the generalization to dimension n, see [**137**, p. 77].) Cactoids have been extensively used. For some other early results, see [**80**, **25**].

A metric continuum X is called *unicoherent* provided that however X be

expressed as the union of two continua M and N, the intersection $M \cap N$ is a continuum. Kuratowski showed [**49**] that the property of unicoherence (for a Peano continuum X) is both cyclicly extensible and reducible. In the same paper he showed that the dimension of X (if X is not a dendrite) is the maximum of the dimensions of its cyclic elements. The localization of this theorem at a point also holds.

In 1913 and 1922, Z. Janiszewski and Miss A. Mullikan had already shown that two bounded subcontinua of certain sorts of space C (e.g. E^2) neither of which disconnects C can have their union disconnect C if and only if their intersection fails to be connected. In 1929, Zippin [**152**] used the cyclic element theory to show that the property of C so defined characterizes the cactoids (among Peano continua), and (among unbounded generalized continuous curves) those spaces whose true cyclic elements are cylinder trees. (A *cylinder tree* is the complement on a simple closed surface of a closed totally disconnected point set.) See also [**26**].

The next year Zippin [**153**] used cyclic elements to show that a generalized continuous curve (in this case a locally compact, connected, and locally connected metric space) which satisfies the Jordan curve theorem non-vacuously is also a cylinder tree.

A fixed point theorem of great interest was proved by Borsuk in 1932. He showed that the fixed point property for Peano continua is both cyclicly extensible and cyclicly reducible [**18**]. (The same paper proves some other properties to be cyclicly extensible and reducible, most notably the property of being an absolute retract.)

In 1926, Scherrer had proved that every continuous transformation of a dendrite into itself has a fixed point. The dendritic structure of a Peano continuum X relative to its cyclic elements led Ayres to seek an analogue of this theorem. He showed [**13**] that a homeomorphism f of X into itself has a "fixed element," i.e. a cyclic element C such that $f(C) \subset C$. After generalizing the cyclic element theory somewhat [**40**], Kelley showed in 1939 [**39**] that, in fact, a homeomorphism of any compact continuum into itself leaves invariant some subcontinuum that has no cutpoint. In 1947 he extended some of his and Ayres's results to arbitrary continuous mappings [**41**].

A transformation of a topological space onto itself is called *periodic* iff for some integer n, the n-fold composition of f with itself is the identity transformation. Ayres studied periodic transformations (and three related "near-periodic" concepts) in connection with the cyclic elements of a Peano continuum in [**14**]. This attack was continued by Whyburn [**137**] and by Schweigert [**84, 85, 86**].

In 1935, C. B. Morrey, Jr. published [**66**] a remarkable application of the theory of cyclic elements to that of path surfaces. Imprecisely, a path surface is any continuous function f defined on a closed disk D into Euclidean 3-dimensional space E^3. Using Moore's notion of upper semi-continuous decomposition, f can be uniquely "factored" into (i.e. represented as the composite of) a continuous mapping $m: D \to M$ and a continuous mapping $l: M \to E^3$ where M, called

the "middle space," is a suitable Peano continuum (a cactoid except that one true cyclic element may be a disk), M is monotone, and l is light. (A mapping is called *light* iff the inverse image of each point is totally disconnected, i.e. contains no connected set with more than one point.)

Morrey was able to reduce many questions in surface theory to corresponding questions about the middle space M and the light map l. In particular, in subsequent papers [**67**, **68**] he was able to use this reduction to express the area of a surface as a sum of areas corresponding to the true cyclic elements of M. Despite a serious error in Morrey's paper (repaired by Youngs in 1944 [**148**]), this reduction became fundamental in area theory. For more detailed discussions, and extensions, see, e.g., [**21**, **60**, **61**, **75**, **147**, **148**, **149**, **150**].

For some other applications of cyclic elements, see [1, 5, 9, 12, 17, 19, 23, 24a, 24b, 29, 31, 34, 35, 54, 62, 74a, 76, 79, 83, 90, 91, 97, 98, 99, 106, 112, 113, 114, 115, 122, 123, 124, 127, 128, 129, 130, 131, 133, 134, 136, 139, 140, 144, 151].

The tremendous barrage of applications of cyclic elements had several effects on the subsequent development of the theory. To start with, it was clearly desirable to obtain a simpler exposition of the principal theorems and their proofs. Whyburn's first exposition [**102**] had based the entire development on the cyclic connectivity theorem. The proof of this was complicated, even in the plane, and considerably more so in more general spaces. So Whyburn thought it desirable to obtain the cyclic element theory results independently of this theorem. He succeeded in [**116**] not only in circumventing the cyclic connectivity theorem, but in extending many of the results to spaces which are connected and locally connected, but not necessarily compact. Since maximal cyclic curves of such spaces might not exist, the key really lies in generalizing the definition of cyclic element. The notion of maximal cyclic curve is replaced by that of set maximal with respect to being connected and having no cutpoint. (Since the resulting "element" need not be cyclic, it was at first called a *nodular set* rather than (as in more recent usage) a *cyclic element*). Cutpoints and endpoints were, of course, retained as cyclic elements.

Whyburn generalized the spaces considered to sets which are merely connected and semi-locally connected in 1939 [**135**].

It is worth noticing that Whyburn's circumvention of the cyclic connectivity theorem enabled him to produce a greatly simplified proof of it by using some elementary cyclic element theory [**119**]. Ayres finally succeeded in producing a very simple proof independent of cyclic element theory [**16**].

An even simpler means of developing cyclic element theory was used by Kuratowski and Whyburn in 1930 [**51**]. (They attribute the idea to Moore [**51**, p. 306 footnote 6].) A point p is said to *separate* two points a and b in a connected set M if $M-\{p\}$ is the union of two mutually separated sets one of which contains a and the other b. Two points a and b are said to be *conjugate* in M provided that no point of M separates a from b in M. Then a cyclic element of a Peano continuum X can be defined as either a cutpoint of X or a set consisting of a non-cutpoint p of X and all the points q of X that are conjugate to p. De-

generate sets of the latter type were shown to be endpoints. For Peano continua the entire theory of cyclic elements was built on this basis, and this paper generally replaced [**106**] as the standard reference on the subject. Kuratowski and Whyburn characterized true cyclic elements in still another fashion (cyclic chain from one of two conjugate points to the other), proved several new properties of cyclic elements, cyclic chains, A-sets, and H-sets, listed fourteen different cyclicly extensible or reducible properties and a criterion by which cyclicly extensible properties can be "mass-produced," listed most of the known applications and several new ones, and gave a complete bibliography of the subject up to 1930.

In addition this paper presented the very useful idea of approximation to a space by means of cyclic chains. The essence of this idea is that there always exists (for a Peano continuum X, at least) a sequence of cyclic chains with diameters tending to zero such that each element of the sequence intersects the union of the previous elements in a single point, but such that the diameters of the components of the complement of the union of the first n chains tend to zero. The union of all the cyclic chains in the sequence covers all of X except possibly for some endpoints. (See also [**137**] where the requirement of local connectedness of X is weakened.)

The concept of *simple link* invented by R. L. Moore coincides with that of true cyclic element for semi-locally connected continua (including Peano continua) but not necessarily for more general spaces. Expositions can be found in [**65**] and [**137**]. Relevant also is [**135**].

Decompositions analogous to the cyclic element decomposition but not strictly refining nor extending the cyclic element theory proper were made by Whyburn in [**122**].

Some further generalizations of the spaces under consideration were made by Harry [**36**], Jones [**37**], Kelley [**40**], and Wallace [**93, 98**].

Much of the "classical" theory of cyclic elements of Peano continua is presented in each of [**33, 50, 59, 65, 75, 137** and **143**].

In 1932, Wilder [**141**] suggested refining cyclic element theory by using techniques from algebraic topology. In a very loose sense, he began this process himself in [**141**] and [**142**] but the development really rests on two papers published by Whyburn in 1934: [**125**], which, among more general results, establishes the existence of the higher cyclic elements, and [**126**] which actually defines these higher cyclic elements and develops the theory. (The treatises mentioned previously as exposing the "classical" theory do no more than hint at this theory. However, it was outlined by Whyburn in [**132**] in a form designed for nonspecialists.)

An essential idea of this development is the use of cuttings other than cutpoints. (By a *cutting* M of a connected topological space X we mean a subset of X whose complement is not connected.) More specifically, a T_r-set (of a compact subset X of Euclidean n-space) is defined to be a closed point set which carries no (essential) r-dimensional cycle. Thus, e.g., a T_0-set is a single point.

The rth order cyclic elements of M or E_r-sets of M are then nondegenerate subsets of M maximal with respect to having no T_r-set as cutting. Thus, e.g., the E_0 sets of M are the true cyclic elements. Some "sample" theorems are:

Two E_r sets of M intersect (if at all) in a T_r set of M.

The dimension of any E_r-set of M is at least $r+1$.

For each r and each E_r in M, there exists a finite sequence of sets $E_0, E_1, \cdots, E_{r-1}$ such that $E_0 \supset E_1 \supset E_2 \supset \cdots \supset E_{r-1} \supset E_r$.

The r-th connectivity number of M is the sum of the r-th connectivity numbers of the E_{r-1}-sets of M.

In order that M should separate R^n it is necessary and sufficient that some E_{n-2} in M should separate R^n.

A considerable portion of the "classical" cyclic element theory was generalized in this manner, from A-sets to extensible and reducible properties.

Subsequent papers developing, extending, and applying the theory of E_r-sets include [**24**, **74**, **81**, and **87**].

A different way of refining cyclic element theory was used by Wallace [**95**, **96**]. Though this theory also results in an analysis that features the dimension of the "cyclic elements" obtained, the technique is based on mappings into spheres.

Several others have also found ways to refine the cyclic element decomposition. Since the "nonseparable components" of a linear graph correspond to cyclic elements of a topological realization of the graph, Whitney's theorem [**100**], that a graph is planar if and only if each nonseparable component is, concerns cyclic reducibility and extensibility of imbeddability in the plane. MacLane [**55**] devised a cyclic element-like decomposition of cyclic elements of a graph based on cuttings by pairs of points to improve Whitney's results. (See also [**56**].)

In 1940, Hall [**30**] produced a more general decomposition of cyclic elements of a Peano continuum, also based on cuttings by pairs of points. Hall's *secondary elements* are built on a notion of *biconjugacy*. Two points a and b were said to be *biconjugate* in cyclicly connected Peano continuum X if every pair of points that separates a from b lies on some free arc of X. The secondary element analogue of a true cyclic element is a suitably chosen point a together with all points biconjugate to a. (The "suitable choosing" consists in picking a so that any point b such that $\{a, b\}$ separates M must lie together with a on some free arc of M.) The applications are concerned chiefly with "secondarily" extensible and reducible properties, e.g. hereditary local connectedness, (Menger) regularity, and (Menger) rationality.

The same year, J. W. T. Youngs [**146**] produced and developed a decomposition somewhat similar to Hall's. For Youngs, point a is *bi-conjugate* to

point b (in a cyclicly connected Peano continuum M) provided that however two points x_1 and x_2 different from a and b be chosen in M, it is true that a and b lie in the same component of $M-\{x_1, x_2\}$. (We obtain "k-conjugacy" by choosing $x_1, x_2, \cdots, x_k$ all different from a and b.) If each of three points (k points) is bi-conjugate (k-conjugate) to each of the others, then the set of all points bi-conjugate (k-conjugate) to all 3 (all k) of them is called a bi-cyclic (k-cyclic) element of M.

In [**138**], Whyburn also generalized the notion of conjugacy, though no parallel with the rest of cyclic element theory is pursued.

Cesari [**20, 21, 22**] formed *fine-cyclic elements* refining the Morrey cyclic element decomposition of surfaces. The idea was taken up and extended by Neugebauer who produced an extensive theory for Peano continua of finite degree of multicoherence [**52, 53, 69, 70, 71, 72**]. (For definition of degree of multicoherence, see e.g., [**137**].) In the Cesari-Neugebauer theory, a B-set (analogue of A-set) is first defined to be a nondegenerate continuum the frontier of each component of whose complement is finite. A *fine-cyclic element* is then a B-set which has no finite cutting. When the space under consideration is unicoherent, the fine-cyclic elements coincide with the Whyburn true cyclic elements.

Others who have produced refinements of the cyclic element decomposition include Kosiński [**45, 46, 47**], McAuley [**58**], and Remage [**78**]. Remage's paper is also concerned with unifying some of the extensions of cyclic element theory.

An interesting by-product of the cyclic element theory is the theory of *acyclic elements* of a Peano continuum X begun by Wallace in 1942. These are simply the components of the set of cutpoints and endpoints of X. The properties of acyclic elements are (roughly) dual to those of cyclic elements. Only a few such properties have been published [**92**].

Another side-benefit is provided by Radó and Reichelderfer's concept of cyclic transitivity [**77**]. They noticed that the relation of conjugacy narrowly fails to be an equivalence relation. Conjugacy is symmetric and reflexive, but not quite transitive, since if b is a cutpoint, we may have a conjugate to b and b conjugate to c without having a conjugate to c. But conjugacy is *cyclicly transitive.* This means that if $a_1, a_2, \cdots, a_n$ is a finite sequence of points, each conjugate to the next, and if a_n is conjugate to a_1, then each of the points is conjugate to each of the others. Abstracting this notion, they produced a "cyclic element" theory based on an arbitrary reflexive, symmetric, cyclicly transitive binary relation. A good many of the classical theorems hold even in this abstract context, which may therefore be thought of as representing the ultimate generalization in the direction of weakening conditions on the space.

This paper grew from reading related to research on a thesis problem suggested by my major professor, Dr. Louis F. McAuley, at the University of Wisconsin. The work has been partly supported by a fellowship in the summer 1963 Research Participation for College Teachers Program at the University of Oklahoma and by an N.S.F. extension grant, GE2687.

References

1. Beatrice Aitchison, Concerning the mapping of locally connected continua onto simple arcs, Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie, Class III, 27 (1934) 130–146.

2. G. E. Albert and J. W. T. Youngs, The structure of locally connected topological spaces, Trans. Amer. Math. Soc., 51 (1942) 637–654.

3. W. L. Ayres, Concerning continuous curves and correspondences, Ann. of Math., 28 (1927) 396–418.

4. ———, On the structure of a plane continuous curve, Proc. Nat. Acad. Sci. U.S.A., 13 (1927) 749–754.

5. ———, Concerning the arc-curves and basic sets of a continuous curve, Trans. Amer. Math. Soc., 30 (1928) 567–578.

6. ———, Continuous curves which are cyclically connected, Bulletin International de l'Académie Polonaise des Sciences et des Lettres, Classe des Sciences Mathématiques Naturelles, (A) (1928) 127–142. (Abstracted: Bull. Amer. Math. Soc., 34 (1928) 139.)

7. ———, Concerning continuous curves in metric space, Amer. J. Math., 51 (1929) 577–594.

8. ———, Concerning the arc-curves and basic sets of a continuous curve. Second paper, Trans. Amer. Math. Soc., 31 (1929) 595–612.

9. ———, Conditions under which every arc of a continuous curve is a subset of a maximal arc of the curve, Math. Ann., 101 (1929) 194–209.

10. ———, Continuous curves homeomorphic with the boundary of a plane domain, Fund. Math., 14 (1929) 92–95.

11. ———, On continua which are disconnected by the omission of any point and some related problems, Monatsh. für Math. und Phys., 36 (1929) 135–148.

12. ———, On the density of the cut points and end points of a continuum, Bull. Amer. Math. Soc., 36 (1930) 659–667.

13. ———, Some generalizations of the Scherrer fixed point theorem, Fund. Math., 16 (1930) 332–336.

14. ———, On transformations having periodic properties, Fund. Math., 33 (1939) 95–103.

15. ———, A note on the definition of arc-sets, Bull. Amer. Math. Soc., 46 (1940) 794–796.

16. ———, A new proof of the cyclic connectivity theorem, Bull. Amer. Math. Soc., 48 (1942) 627–629.

17. R. H. Bing, Some monotone decompositions of a cube, Ann. of Math., (2) 61 (1955) 279–288.

18. K. Borsuk, Einige Sätze über stetige Streckenbilder, Fund. Math., 18 (1932) 198–213.

19. ———, Über die Abbildung der metrischen kompakten Räume auf die Kreislinie, Fund. Math., 20 (1933) 224–231.

20. L. Cesari, Fine-cyclic elements of surfaces of the type ν, Riv. Mat. Univ. Parma, 7 (1956) 149–185.

21. ———, Surface area, Ann. of Math. Studies 35, Princeton University Press, 1956.

22. ———, A new process of retraction and the definition of fine cyclic elements, An. Acad. Brasil. Ci., 29 (1957) 1–7.

22a. ——— and C. J. Neugebauer, On the coincidence of Geöcze and Lebesgue areas, Duke Math. J., 26 (1959) 147–153.

23. S. Claytor, Topological immersion of Peanian continua in a spherical surface, Ann. of Math., 35 (1934) 809–835.

24. J. Gary, Higher cyclic elements, Pacific J. Math., 9 (1959) 1061–1070.

24a. W. H. Fleming, Nondegenerate surfaces of finite topological type, Trans. Amer. Math. Soc., 90 (1959) 323–335.

24b. ———, Nondegenerate surfaces and fine-cyclic surfaces, Duke Math. J., 26 (1959) 137–146.

25. H. M. Gehman, A special type of upper semi-continuous collection, Proc. Nat. Acad. Sci. U.S.A., 16 (1930) 609–613.

26. A. Grzegorczyk and C. Kuratowski, On Janiszewski's property of topological spaces, Ann. Soc. Polon. Math., 25 (1953) 169–182 (1952).

27. H. Hahn, Mengentheoretische Charakterisierung der stetigen Kurve, Sitzungsber, Akad. Wiss. Wien, Math.-Nat. Klasse, 123 (1914) 2433–2489.

28. T. Hailperin, On contiguous point spaces, Bull. Amer. Math. Soc., 35 (1939) 172–174.

29. D. W. Hall, On the nonalternating images of linear graphs, Amer. J. Math., 59 (1937) 575–584.

30. ———, On a decomposition of true cyclic elements, Trans. Amer. Math. Soc., 47 (1940) 305–322.

31. —— and W. T. Puckett, Jr., Conditions for the continuity of arc-preserving transformations, Bull. Amer. Math. Soc., 47 (1941) 468–475.

32. ——— and ———, Strongly arc-wise connected spaces, Amer. J. Math., 63 (1941) 554–562.

33. ——— and G. L. Spencer II, Elementary topology, Wiley, New York, 1955.

34. ——— and G. T. Whyburn, Arc- and tree-preserving transformations, Trans. Amer. Math. Soc., 48 (1940) 63–71.

35. ——— and J. W. T. Youngs, Comments on the cores of certain classes of spaces, Ann. of Math., 48 (1947) 710–716.

36. C. H. Harry, Concerning the geometry of acyclic sets, Amer. J. Math., 56 (1934) 233–253.

37. F. B. Jones, Almost cyclic elements and simple links of a continuous curve, Bull. Amer. Math. Soc., 46 (1940) 775–783.

38. C. Jordan, Cours d'Analyse, 2me éd., Vol. 1, Gauthier-Villars, Paris, 1893.

39. J. L. Kelley, Fixed sets under homeomorphisms, Duke Math. J., 5 (1939) 535–537.

40. ———, A decomposition of compact continua and related theorems on fixed sets under continuous transformations, Proc. Nat. Acad. Sci., 26 (1940) 192–194.

41. ———, Simple links and fixed sets under continuous mappings, Amer. J. Math., 69 (1947) 348–356.

42. ———, General topology, Van Nostrand, Princeton, N. J., 1955.

43. E. C. Klipple, Two-dimensional spaces in which there exist contiguous points, Trans. Amer. Math. Soc., 44 (1938) 250–276.

44. D. König, Theorie der endlichen und unendlichen Graphen, Chelsea, New York, 1950, Chapter 17.

45. A. Kosiński, On some problems connected with the topology of manifolds, Proc. of the Intern. Congress of Math., 14–21 August 1958, pp. 427–432.

46. ———, Some theorems about two-dimensional polyhedra, Fund. Math., 47 (1959) 1–28.

47. ———, On strongly cyclic elements (not published).

48. I. S. Krule, Structs on the 1-sphere, Duke Math. J., 24 (1957) 405–413.

49. C. Kuratowski, Quelques applications d'éléments cycliques de M. Whyburn, Fund. Math., 14 (1929) 138–144.

50. ———, Topologie II, Polska Akademia Nauk Monografie Matematyczne 21, Warsaw, 1961, Hafner, New York.

51. ——— and G. T. Whyburn, Sur les éléments cycliques et leurs applications, Fund. Math., 16 (1930) 305–331.

52. A. Lauria and C. J. Neugebauer, A characterization of multicoherent Peano spaces, dittoed notes, Purdue University, 1959.

53. ——— and ———, Retraction onto dendrites, dittoed notes, Purdue University, 1959.

54. M. B. Lieberknecht, Some characterizations of unicoherent Peano spaces, master's thesis, Ohio State University, 1950.

55. S. MacLane, A structural characterization of plane combinatorial graphs, Duke Math. J., 3 (1937) 460–472.

56. ——— and V. W. Adkisson, Extensions of the homeomorphisms on the sphere, in Lectures in Topology, the University of Michigan conference of 1940, U. Mich. Press, 1941.

57. S. Mazurkiewicz, Sur les lignes de Jordan, Fund. Math., 1 (1920) 166–209.

58. L. F. McAuley, Local cyclic connectedness, I. Cyclic subelements of cyclic elements, Notices, Amer. Math. Soc., 5 (1958) 698 (abstract).

59. K. Menger, Kurventheorie, Teubner, Leipzig, 1932, chapter 12.

60. E. J. Mickle and C. J. Neugebauer, Weak and strong cyclic additivity, Riv. Mat. Univ. Parma, 7 (1956) 243–253.

61. ——— and T. Radó, On cyclic additivity theorems, Trans. Amer. Math. Soc., 66 (1949) 347–365.

62. D. D. Miller, Extension and reduction for certain types of continuous transformations, Amer. J. Math., 64 (1942) 215–218.

63. R. L. Moore, Concerning upper semi-continuous collections, Monatsh. für Math. und Phys., 36 (1929) 81–88.

64. ———, Fundamental theorems concerning point sets, The Rice Institute Pamphlet, 23 (1936) 1–74.

65. ———, Foundations of point set theory, revised ed., Amer. Math. Soc. Colloq. Publ. 13, 1963.

66. C. B. Morrey, Jr., The topology of (path) surfaces, Amer. J. Math., 57 (1935) 17–50.

67. ———, An analytic characterization of surfaces of finite Lebesgue area, Part I, Amer. J. Math., 57 (1935) 692–702.

68. ———, An analytic characterization of surfaces of finite Lebesgue area, Part II, Amer. J. Math., 58 (1936) 313–322.

69. C. J. Neugebauer, A decomposition of a Peano space, dittoed notes, Purdue University, 1957.

70. ———, A fine-cyclic additivity theorem for a functional, Illinois J. Math., 2 (1958) 121–136.

71. ———, B-sets and fine-cyclic elements, Trans. Amer. Math. Soc., 88 (1958) 121–136.

72. ———, Local A-sets, B-sets, and retractions, Illinois J. Math., 2 (1958) 386–395.

73. G. Peano, Sur une courbe qui remplit toute une aire plane, Math Ann., 36 (1890) 157–160.

74. W. T. Puckett, Jr., A Theorem concerning homologies in a compact space, Fund. Math., 31 (1938) 22–26.

75. T. Radó, Length and area, Amer. Math. Soc. Colloq. Publ. 30, 1948.

76. ———, On continuous mappings of Peano spaces, Trans. Amer. Math. Soc., 58 (1945) 420–454.

77. ——— and P. Reichelderfer, Cyclic transitivity, Duke Math. J., 6 (1940) 474–485.

78. R. Remage, On cuttings of finite sets of points, Duke Math. J., 31 (1964) 359–370.

79. ———, Invariance and periodicity properties of nonalternating in the large transformations, dissertation, U. of Pennsylvania, 1950.

79a. ———, An extension of a fine-cyclic additivity theorem for a functional, Illinois J. of Math., 7 (1963) 548–577.

80. J. H. Roberts, A note concerning cactoids, Bull. Amer. Math. Soc., 36 (1930) 894–896.

81. ——— and N. E. Steenrod, Monotone transformations of two-dimensional manifolds, Ann. of Math., 39 (1938) 851–862.

82. N. E. Rutt, Concerning the cut points of a continuous curve when the arc curve AB contains exactly N independent arcs, Amer. J. Math., 51 (1929) 217–246.

83. G. E. Schweigert, Concerning nonalternating interior transformations, Bull. Amer. Math. Soc., 44 (1938) 636 (abstract).

84. ———, Fixed elements and periodic types for homeomorphisms on semi-locally connected continua, Proc. Nat. Acad. Sci. U.S.A., 29 (1943) 52–54.

85. ———, Minimal A-sets, infinite orbits, and fixed elements, Bull. Amer. Math. Soc., 49 (1943) 754–758.

86. ———, Fixed elements and periodic types for homeomorphisms on semi-locally connected continua, Amer. J. Math., 66 (1944) 229–244.

87. Arthur B. Simon, n-Cyclic elements I, Duke Math. J., 24 (1957) 1–7.

88. N. E. Steenrod, Characterizations of certain finite curve sums, Amer. J. Math., 56 (1934) 558–568.

89. ———, Finite arc-sums, Fund. Math., 23 (1934) 38–53.

90. A. D. Wallace, On 0-regular transformations, Amer. J. Math., 62 (1940) 277–284.

91. ———, Concerning relatively nonalternating transformations, Proc. Nat. Acad. Sci. U.S.A., 27 (1941) 182–185.

92. ———, The acyclic elements of a Peano space, Bull. Amer. Math. Soc., 47 (1941) 778–780.

93. ———, Monotone transformations, Duke Math. J., 9 (1942) 487–506.

94. ———, Reducible properties of Peano spaces, Anais Acad. Brasil. Sci., 14 (1942) 47–49.

95. ———, Generalized arc-sets, Proc. Nat. Acad. Sci. U.S.A., 31 (1945) 414–417.

96. ———, Extension sets, Trans. Amer. Math. Soc., 59 (1946) 1–13.

97. ———, Cyclic invariance under multi-valued maps, Bull. Amer. Math. Soc., 55 (1949) 820–824.

98. ———, Endelements and the inversion of continua, Duke Math. J., 16 (1949) 141–144.

99. C. H. Wheeler, III, A type of homogeneity for continuous curves, Amer. J. Math., 59 (1937) 315–326.

100. H. Whitney, Nonseparable and planar graphs, Trans. Amer. Math. Soc., 34 (1932) 339–362.

101. G. T. Whyburn, Concerning continua in the plane, Trans. Amer. Math. Soc., 29 (1927) 369–400.

102. ———, Cyclicly connected continuous curves, Proc. Nat. Acad. Sci. U.S.A., 13 (1927) 31–38.

103. ———, Some properties of continuous curves, Bull. Amer. Math. Soc., 33 (1927) 305–308.

104. ———, Concerning certain kinds of continuous curves and other continua, Bull. Amer. Math. Soc., 33 (1927) 657 (abstract).

105. ——, Concerning the open sets of a plane continuous curve, Proc. Nat. Acad. Sci. U.S.A., 13 (1927) 650–657.

106. ———, Concerning the structure of a continuous curve, Amer. J. Math., 50 (1928) 167–194.

107. ———, Concerning Menger regular curves, Fund. Math., 12 (1928) 264–294.

108. ———, Concerning the cyclic elements of continuous curves, Bull. Amer. Math. Soc., 34 (1928) 700 (abstract).

109. ———, On a problem of W. L. Ayres, Fund Math., 11 (1928) 296–301.

110. ———, Continuous curves and arc sums, Fund. Math., 14 (1929) 103–106.

111. ———, On regular points of continua and regular curves of at most order n, Bull. Amer. Math. Soc., 35 (1929) 218–224.

112. ———, On the structure of plane closed point sets which are accessible from certain subsets of their complements, Proc. Nat. Acad. Sci. U.S.A., 14 (1929) 657–666.

113. ———, Concerning continuous curves without local separating points, Bull. Amer. Math. Soc., 36 (1930) 631–632 (abstract).

114. ———, Concerning points of continua defined by certain im kleinen properties, Math. Ann., 102 (1930) 313–336.

115. ———, On the set of all cut points of a continuous curve, Fund. Math., 15 (1930) 185–194.

116. ———, On the structure of connected and connected im kleinen point sets, Trans. Amer. Math. Soc., 32 (1930) 926–943.

117. ———, The rationality of certain continuous curves, Bull. Amer. Math. Soc., 36 (1930) 522–524.

118. ———, On the cyclic and higher connectivity of locally connected spaces, Amer. J. Math., 53 (1931) 427–442.

119. ———, On the cyclic connectivity theorem, Bull. Amer. Math. Soc., 37 (1931) 429–433.

120. ———, On the relative cyclic element decomposition of sub-continua of a continuous curve, Bull. Amer. Math. Soc., 37 (1931) 39 (abstract).

121. ———, On the decomposability of closed sets into a countable number of simple sets of various types, Amer. J. Math., 54 (1932) 169–175.

122. ———, Decompositions of continua by means of local separating points, Amer. J. Math., 55 (1933) 437–457.

123. ———, On the existence of totally imperfect and punctiform connected subsets in a given continuum, Amer. J. Math., 55 (1933) 146–152.

124. ———, Sets of local separating points of a continuum, Bull. Amer. Math. Soc., 39 (1933) 97–100.

125. ———, Concerning maximal sets, Bull. Amer. Math. Soc., 40 (1934) 159–164.

126. ———, Cyclic elements of higher orders, Amer. J. Math., 56 (1934) 133–146.

127. ———, Nonalternating transformations, Amer. J. Math., 56 (1934) 294–302.

128. ———, Concerning continua of finite degree and local separating points, Amer. J. Math., 57 (1935) 11–16.

129. ———, On sequences and limiting sets, Fund. Math., 25 (1935) 408–426.

130. ———, Regular convergence and monotone transformations, Amer. J. Math., 57 (1935) 902–906.

131. ———, On continua of condensation, Amer. J. Math., 58 (1936) 705–708.

132. ———, On the structure of continua, Bull. Amer. Math. Soc., 42 (1936) 49–73.

133. ———, The mapping of Betti groups under interior transformations, Duke Math. J., 4 (1938) 1–8.

134. ———, Nonalternating interior retracting transformations, Ann. of Math., (2) 40 (1939) 914–921.

135. ———, Semi-locally connected sets, Amer. J. Math., 61 (1939) 733–749.

136. ———, The existence of certain transformations, Duke Math. J., 5 (1939) 647–655.

137. ———, Analytic topology, Amer. Math. Soc. Colloq. Pub., 28 (1942).

138. ———, Coherent and saturated collections, Trans. Amer. Math. Soc., 57 (1945) 287–298.

139. ———, Uniqueness of the inverse of a transformation, Duke Math. J., 12 (1945) 317–323.

140. ———, On monotone retractability into simple arcs, Bull. Amer. Math. Soc., 52 (1946) 109–112.

141. R. L. Wilder, Point sets in three and higher dimensions and their investigation by means of a unified analysis situs, Bull. Amer. Math. Soc., 38 (1932) 649–692.

142. ———, On the linking of Jordan continua in E_n by $(n-2)$-cycles, Ann. of Math., 34 (1933) 441–449.

143. ———, Topology of Manifolds, Amer. Math. Soc. Colloq. Publ. 32, 1949.

144. G. S. Young, Jr., On continua whose links are nonintersecting, Bull. Amer. Math. Soc., 50 (1944) 920–925.

145. J. W. T. Youngs, Arc-spaces, Duke Math. J., 7 (1940) 68–84.

146. ———, k-Cyclic elements, Amer. J. Math., 62 (1940) 449–456.

147. ———, Curves and surfaces, this MONTHLY, 51 (1944) 1–11.

148. ———, The topological theory of Fréchet surfaces, Ann. of Math., (2) 45 (1944) 753–785.

149. ———, Remarks on cyclic additivity, Bull. Amer. Math. Soc., 55 (1949) 427–432.

150. ———, The representation problem for Fréchet surfaces, Memoirs of the Amer. Math. Soc. 8, 1951.

151. K. Zarankiewicz, On the category of the set of cut points of continua of a certain type, Czechoslovak Math. J., 1 (76) (1951) 57–62.

152. L. Zippin, A study of continuous curves and their relation to the Janiszewski-Mullikan theorem, Trans. Amer. Math. Soc., 31 (1929) 744–770.

153. ———, On continuous curves and the Jordan curve theorem, Amer. J. Math., 52 (1930) 331–350.

A FIXED POINT THEOREM FOR CONNECTED MULTI-VALUED FUNCTIONS

RAYMOND E. SMITHSON, U. S. Naval Ordnance Test Station
China Lake, California

1. Introduction. Let X and Y be Hausdorff spaces. A multi-valued function on X into Y is a point to set correspondence. Let $F: X \to Y$ denote a multi-valued function. If $A \subset X$, then the image of A is $F(A) = \bigcup \{F(x): x \in A\}$ and if $B \subset Y$, then the inverse of B is $F^{-1}(B) = \{x \in X: F(x) \cap B \neq \emptyset\}$ where $\emptyset$ denotes the empty set. In particular $F^{-1}(y) = \{x: y \in F(x)\}$. If $F: X \to X$, then $x \in X$ is a fixed point for the multi-valued function F in case $x \in F(x)$. In the sequel the term function will mean multi-valued function.

The concept of continuity has been defined for multi-valued functions in several different ways, and W. Strother [**10**] has studied the relationships that exist between the various definitions of continuity. C. Berge discusses multi-valued functions in [**3**], and develops many of the fundamental properties. Many authors have obtained fixed point theorems for functions that satisfy one or more of the various definitions of continuity; see for example, Begle [**2**], Fuller [**5**], Strother [**11**], Wallace [**12, 13**], Ward [**16, 17, 18**], and their references. On the other hand, Hamilton [**6**] has obtained a fixed point theorem without the assumption of continuity. The purpose of this note is to prove a fixed point theorem for a class of functions which are not necessarily continuous.

A multi-valued function F is called point-closed if $F(x)$ is a closed set for each $x \in X$. If $F(C)$ is a connected subset of Y whenever C is a connected subset of X, then F is called a connected function. If the inverse under F of a connected set is connected, then F is called inverse-connected. If F is both connected and inverse-connected, call F biconnected.

Our main result is Theorem 1 below which states: A biconnected, point-closed function on a tree into itself has a fixed point. This theorem extends results of Wallace [**12**], Capel and Strother [**1**], and Ward [**17, 18**], to functions which need not be continuous.

2. Trees. The purpose of this section is to define a tree, and to exhibit the principal results used in the sequel. In particular, the order theoretic properties of trees are developed and used to derive the fundamental lemmas required in the proof of Theorem 1. The idea of introducing an order and a topology independently to study properties of spaces was due to A. D. Wallace. L. E. Ward [**14–18**] has studied the order theoretic properties of topological spaces, and has used these properties to prove fixed point theorems. Also, G. D. Birkhoff [**4**] and L. Nachbin [**7, 8**] have considered orders and independently obtained part of Lemma 2.

DEFINITION. *A point $z \in X$ is said to separate the points x, $y \in X$ if and only if there exist sets A, $B \subset X$ with $x \in A$, $y \in B$ such that $\overline{A} \cap B = A \cap \overline{B} = \emptyset$ and such that $X - \{z\} = A \cup B$ (where $\overline{A} = Cl(A)$ denotes the closure of A). We say that x and y are separated by z in X.*

A *tree* is a compact, connected, Hausdorff space X in which any two points can be separated in X by a third point.

REMARK. Ward [**14, 15**] has shown that a tree is locally connected.

Let X be a tree and let $e \in X$. Define $x \leqq y$ if and only if $x=e$, $x=y$, or x separates e and y in X; the element e is called the least element of X. Define $M(x)=\{a: x \leqq a\}$ and $L(x)=\{a: a \leqq x\}$. Ward [**14**] has shown that $\leqq$ is a partial order on the set X. In the following X will be a tree with least element e and the above partial order.

Lemma 1 is from [**14**] and [**15**].

LEMMA 1. *The partial order $\leqq$ satisfies:*

(i) *$\leqq$ is semi-continuous; i.e., $L(x)$ and $M(x)$ are closed sets for each $x \in X$.*

(ii) *$\leqq$ is order dense.*

(iii) *If $x \in X$ and $y \in X$, then $L(x) \cap L(y)$ is a nonempty chain,*

(iv) *If $x \in X$, the $M(x)-x$ is open,*

where a chain is a linearly ordered subset of X.

If $x \leqq y$, define $[x, y]=\{z: x \leqq z \leqq y\}$ and define $(x, y)=\{z: x<z<y\}$, where $x<z$ means $x \leqq z$ but $x \neq z$. One defines $[x, y)$ and $(x, y]$ in a similar manner. Note that $[x, x]=\{x\}$.

The first part of Lemma 2 is due to A. D. Wallace [**13**].

LEMMA 2. *Each compact, connected subset of X contains a least element, and every closed chain contains a maximal element. Further, if $x \leqq y$, $[x, y]$ is a compact, connected chain.*

Proof. The first statement follows from Theorem 1 of [**14**]. The second statement follows from Lemma 1 and the first statement.

LEMMA 3. *If $\mathfrak{N}$ is a nest of sets, of the form $[e, x]$, in X, then there exists a point x_0 such that $[e, x] \subset [e, x_0]$ for all $[e, x] \in \mathfrak{N}$ and if $[e, x] \subset [e, x_0)$, then $[e,x]$ is included in a member of $\mathfrak{N}$. From this last it follows that if $e \leqq x < x_0$, then $x \in \bigcup \mathfrak{N}$.*

Proof. The set $\bigcup \mathfrak{N}$ is connected as each $N \in \mathfrak{N}$ is connected and contains e. Thus, $\mathrm{Cl}(\bigcup \mathfrak{N})$ is compact and connected, and hence contains a maximal element x_0 by Lemma 2. Then Lemma 1 together with the fact that $\mathfrak{N}$ is a nest of sets of the form $[e, x]$ show that x_0 is the desired point.

REMARK. Let $x \in X$. Then $X-\{x\}=X-M(x) \cup M(x)-\{x\}$. By Lemma 1, $M(x)$ is closed and $M(x)-\{x\}$ is open. Further, $X-M(x)$ and $M(x)-\{x\}$ are disjoint. Thus, the uniands $X-M(x)$ and $M(x)-\{x\}$ are open and disjoint.

LEMMA 4. *If $x_0 \notin A$, $A \cap M(x_0) \neq \varnothing$, and if there is an $x \in A$ such that $x_0 \nleqq x$, then x_0 separates A in X. Thus, A is not connected.*

Proof. Let $A_1=\{x \in A: x_0 \nleqq x\}$ and let $A_2=A \cap M(x_0)$. Since $x_0 \notin A$, $M(x_0)-x_0$ is an open set containing A_2 which does not meet A_1. Further since $M(x_0)$

is closed, the complement of $M(x_0)$ is an open set containing A_1 which does not meet A_2. Hence the lemma follows.

We shall use the following notation: If $A \subset X$ and $x \leqq y$ $(x<y)$ for all $y \in A$, write $x \leqq A$ $(x<A)$. Also, if $A \subset X$, define $A^{\#}$ by: $A^{\#} = \{x \in X : x \nleqq y \text{ for all } y \in A\}$.

LEMMA 5. *Let $x \in X$. If $F(x)$ is connected and if $x \notin F(x)$, then either $x < F(x)$ or $x \in F(x)^{\#}$.*

Proof. Suppose that $x \nless y$ for some $y \in F(x)$ and suppose that $x \notin F(x)^{\#}$, i.e., there is a $z \in F(x)$ such that $x \leqq z$. The latter supposition implies that $F(x) \cap M(x) \neq \emptyset$. Then since $x \notin F(x)$ and $x \nless y \in F(x)$, Lemma 4 implies that $F(x)$ is not connected, a contradiction. Hence, either $x < F(x)$ or $x \in F(x)^{\#}$.

LEMMA 6. *Let F be a point-closed, connected function on the tree X into itself such that $x \notin F(x)$ for all $x \in X$. Let $x_0 \in X$ and suppose that for $e \leqq x < x_0$, we have $x < F(x)$. Then $x_0 < F(x_0)$.*

Proof. If $x_0 = e$, this is trivial. Thus suppose $x_0 \neq e$. By hypothesis $x_0 \notin F(x_0)$ and $F(x_0)$ is connected. Thus, we suppose that $x_0 \in F(x_0)^{\#}$. Let $A = F(x_0) \cap [e, x_0]$ and let $x_1 = \max (A \cup \{e\})$ (x_1 exists by Lemma 2). Then $e \leqq x_1 < x_0$. Let z be such that $e \leqq x_1 < z < x_0$. Then $z \in F(x_0)^{\#}$ and $z < F(z)$. Hence, z separates the set $F([z, x_0])$, but this is a contradiction, since F is a connected function.

3. The Main Theorem. We now state our main result.

THEOREM 1. *Let X be a tree. If $F: X \rightarrow X$ is a biconnected, point-closed function on X into X, then F has a fixed point.*

Proof. Suppose that F does not have a fixed point. Let X be partially ordered as in Section 2 with minimal element e. Then $e \notin F(e)$. Thus, $x \in F(e)$ implies that $e < x$. Hence $e < F(e)$. Let $\mathcal{S}$ be defined by:

$$\mathcal{S} = \{[e, x] : y < F(y) \text{ for all } y \in [e, x]\}.$$

Since $[e, e] = \{e\}$, $\mathcal{S} \neq \emptyset$. Partial order $\mathcal{S}$ by inclusion. Then, by Lemmas 3 and 6 and Zorn's lemma, $\mathcal{S}$ contains a maximal element. Let $[e, x_0]$ be a maximal element; then $x_0 < F(x_0)$, and if x_1 is the least element of $F(x_0)$, then $x_0 < x_1$. By the maximality of $[e, x_0]$, there exists a y such that $x_0 < y < x_1$ and such that $y \in F(y)^{\#}$ (by Lemma 5). We now assert that $x_0 < x < y$ implies that $x \in F(x)^{\#}$. For suppose that $x_0 < x < y$ and that $x < F(x)$. Let $x_2 = \min F(x)$, and let $A = [x, x_1] \cup [x, x_2]$. Then A is connected and thus, $F^{-1}(A)$ is connected. Since $x_1 \in F(x_0)$ and $x_2 \in F(x)$, both $x_0, x \in F^{-1}(A)$. Further, since $x_0 < x$ and $F^{-1}(A)$ is connected, $[x_0, x] \subset F^{-1}(A)$; for otherwise, by Lemma 4, some point $z \in [x_0, x]$ would separate $F^{-1}(A)$ in X. Consequently, if $z \in [x_0, x)$, $F(z) \cap A \neq \emptyset$, and since $x \leqq A$, and $z \notin F(z)$, Lemma 5 implies that $z < F(z)$ for all $z \in [x_0, x)$. Hence for all $z \in [e, x]$, we have $z < F(x)$. But this contradicts the maximality of $[e, x_0]$. We have shown that if $x_0 < x \leqq y$, then $x \in F(x)^{\#}$. However, $F([x_0, y])$ is connected, since F maps connected sets onto connected sets and $[x_0, y]$ is con-

nected by Lemma 2. On the other hand, $y < F(x_0)$ and so $F([x_0, y])$ meets both $M(y)$ and $X - M(y)$. Further, from the facts that $x \in F(x)^\#$ for all $x_0 < x \leqq y$, and $y < F(x_0)$ we see that $y \notin F([x_0, y])$. Thus, y separates $F([x_0, y])$ in X, a contradiction. Consequently, $x_0 < F(x_0)$ is impossible and so there exists a point $x \in X$ such that $x \in F(x)$.

4. Some Examples. In this section we present some examples which show that the condition that F be biconnected does not imply that F is continuous.

Example 1. Let $X = [0, 1]$ and define F_1 by: $F_1(x) = \frac{1}{2} \cdot x$ if $0 \leqq x < \frac{1}{2}$, $F_1(\frac{1}{2}) = [\frac{1}{4}, \frac{1}{2}]$, $F_1(x) = x$ for $\frac{1}{2} < x \leqq 1$. This function is point closed and biconnected but the inverse of the open set $(\frac{1}{4}, \frac{1}{2})$ is not open. Thus, F_1 is not lower semi-continuous.

Example 2. Let $X = [1, 2]$ and define F_2 by: $F_2(1) = 1$, $F_2(x) = [1, x]$ for $1 < x < 2$ and $F_2(2) = 2$. This function is point closed and biconnected but $F_2^{-1}(x) = [x, 2)$ when $1 \leqq x < 2$. Thus, F_2 is not upper semi-continuous.

Example 3. Let $X = [0, 2]$ and define F_3 by: $F_3(x) = F_1(x)$ if $x \in [0, 1]$, and $F_3(x) = F_2(x)$ if $x \in [1, 2]$. Then F is biconnected and point closed but neither upper nor lower semi-continuous.

Strother [11] has given an example of a continuous, biconnected function on the two cell into itself which does not have a fixed point. In [9] there is an example of a continuous, biconnected function on an arcwise connected, acyclic, metric continuum into itself which does not have a fixed point. This latter space is not locally connected, it is not unicoherent and it contains a nest of arcs which is not contained in an arc. We then conjecture that Theorem 1 can be extended to acyclic, arcwise connected continua in which each nest of arcs is contained in an arc.

Presented to the American Mathematical Society April 18, 1964.

References

1. C. E. Capel and W. L. Strother, Multi-valued functions and partial order, Port. Math., 17 (1958) 41–47.

2. E. G. Begle, A fixed point theorem, Ann. Math., 51 (1950) 544–550.

3. C. Berge, Topological Spaces, Macmillan, New York, 1963.

4. G. Birkhoff, Lattice theory, rev. ed. Amer. Math. Soc. Colloq. Publ., 1948.

5. F. B. Fuller, Fixed points of multiple-valued transformations, Bull. Amer. Math. Soc., 67 (1961) 165–169.

6. O. H. Hamilton, Fixed points for certain noncontinuous transformations, Proc. Amer. Math. Soc., 8 (1957) 750–756.

7. L. Nachbin, Sur les espaces topologiques ordonnés, C. R. Acad. Sci. Paris, 226 (1948) 381–382.

8. ———, Topologia e Ordem, Chicago, 1950.

9. R. E. Smithson, Changes of topology and fixed points for multi-valued functions, Proc. Amer. Math. Soc., to appear.

10. W. L. Strother, Continuous multi-valued functions, Bol. Soc. Mat. São Paulo, 10 (1955) 87–120.

11. ———, Fixed points, fixed sets, and M-retracts, Duke Math. J., 22 (1955) 551–556.

12. A. D. Wallace, A fixed point theorem for trees, Bull. Amer. Math. Soc., 47 (1941) 757–760.
13. ———, A fixed point theorem, Bull. Amer. Math. Soc., 51 (1945) 413–416.
14. L. E. Ward, Jr., Partially ordered topological spaces, Proc. Amer. Math. Soc., 5 (1954) 144–161.
15. ———, A note on dendrites and trees, Proc. Amer. Math. Soc., 5 (1954) 992–994.
16. ———, Mobs, trees and fixed points, Proc. Amer. Math. Soc., 8 (1957) 798–804.
17. ———, A fixed point theorem for multi-valued functions, Pac. J. Math., 8 (1958) 921–927.
18. ———, Characterization of the fixed point property for a class of set-valued mappings, Fund. Math., 2 (1961) 159–164.

ON A NEW PROBLEM IN NUMBER THEORY

ROGER CROCKER, University of Toledo

In this article, the following type of problem is considered. Instead of expressions a^n or n^a, where a is fixed and n may take on different values, one wishes to consider n^n (where both exponent and base change together) and investigate certain problems arising in connection with it. Some of these problems are analogous to those arising with n^a and a^n.

The starting point is the following problem, analogous to the "primitive root and belonging to a given exponent" problem for a^n.

As is seen immediately, $(p-1)^{(p-1)} \equiv 1 \pmod{p}$, where p is an arbitrary odd prime. Also $1^1 \equiv 1 \pmod{p}$. Now for a fixed p does there exist an n, $1<n<p-1$, such that $n^n \equiv 1 \pmod{p}$, or is $p-1$ the lowest nontrivial integer having this property? If $p-1$ is the lowest nontrivial integer having this property, p will be said to be *irreducible*; otherwise, p will be said to be reducible (for n).

From consideration of this problem, one is led to Theorem I, which is derived immediately below and which produces a class of irreducible primes.

Notation. Throughout this paper, p denotes an arbitrary odd prime, or (if specifically stated) an arbitrary prime of a certain form. Throughout the derivation of Theorem I, it is understood that n, $s \neq 1$ or $p-1$.

Let s be any integer having the property, call it the fundamental property, that $s^s \equiv 1 \pmod{p}$, $1<s<p-1$, p fixed. Then s must belong to some exponent $d \pmod{p}$, where d is a divisor of both $p-1$ and of s. (However, $d \neq 1$ or $p-1$ clearly; also $d \neq 2$, as otherwise $s=1$, $p-1$ which is excluded.) $\therefore d \mid s$. (If s did not belong to d, d would not have to divide s. It should further be noted that if $n^d \equiv 1 \pmod{p}$, $1<n<p-1$, $d \mid n$ implies n has the fundamental property, whether or not n belongs to $d \pmod{p}$.)

Now suppose $p-1=2q$, q a prime of the form $4h+3$. The only divisors of $p-1$ are 1, 2, q, $2q$. Hence s must belong to $q \pmod{p}$ and $q=(p-1)/2$ must divide s. Thus $s=(p-1)/2$. But since $(p-1)/2$ is odd, $(p-1)^{(p-1)/2} \equiv -1 \pmod{p}$.

And 2 is a quadratic residue of p as $p=8h+7$; hence $2^{(p-1)/2}\equiv 1 \pmod{p}$. Thus

$$\left(\frac{p-1}{2}\right)^{(p-1)/2} \equiv -1 \pmod{p}.$$

Therefore $(p-1)/2$ does not have the fundamental property, and p is irreducible. It follows that

THEOREM I. *If $p=2q+1$ where q is a prime and of the form $4h+3$, then p is irreducible.*

It is natural to ask whether there is an infinity of irreducible primes. This question, like the corresponding one for primitive roots, cannot yet be definitely answered, though after Theorem I an affirmative answer seems probable.

Remark. For $p=4x+3$, $(p-1)^{(p-1)/2}\equiv -1 \pmod{p}$. Hence a sufficient condition that $p=4x+3$ be reducible is that $2^{(p-1)/2}\equiv -1 \pmod{p}$ and hence that 2 be a quadratic nonresidue of p. Therefore, $p=8x+3$ is reducible. Since for $p=4x+1$, $(p-1)^{(p-1)/2}\equiv 1 \pmod{p}$ instead of $-1 \pmod{p}$, it follows that the sufficient condition for reducibility is that $2^{(p-1)/2}\equiv 1 \pmod{p}$ or that 2 be a quadratic residue of p. Hence $p=8x+1$ is reducible, and it follows that *there is an infinity of reducible primes of either the form $4x+1$ or the form $4x+3$.*

The first paragraph of the proof of Theorem I suggests a method for determining whether p is irreducible (if $p=8x+1$ or $8x+3$, it is known from above to be reducible) which is faster than such crude, obvious methods as trying every n^n (or even n^n for only those n such that $(n, p-1)\neq 1$) especially if each of the prime factors of $p-1$ is raised to a small power. This increase in speed results mainly because the exponents involved in the calculations are, on the average, far smaller, as can be seen by a comparison of the methods.

Method: Start with the highest $d<(p-1)/2$; consider $x^d\equiv 1 \pmod{p}$ for same d. Then for x, try every multiple of that particular d, i.e. every multiple which is >1 and $<p-1$. If one of these multiples satisfies that congruence, p is reducible. If no multiple does, one follows the same procedure for the following (lower) d, and if one arrives at a negative result again, for the d after that, and so on. If one obtains a negative result for every d, p is irreducible, since any n, with $1<n<p-1$, having the fundamental property must belong to some d (and be a solution of $x^d\equiv 1 \pmod{p}$) and hence be a multiple of that d; every possible integer will have been tried and the test is complete.

Actually one may start with any d, but $d=(p-1)/2$ should be excluded as it has been taken into account above, where it is shown that if $p=8x+1$ or if $p=8x+3$, then it is reducible. Starting with the highest $d<(p-1)/2$ may be better in that on the average, reducible primes can be found to be reducible more quickly (i.e. with fewer trials).

A topic suggested by the preceding problem is the study of the residues of $n^n \pmod{p}$, $1\leqq n\leqq p-1$. This study leads to several results. Since

$$a^a(p-a)^{(p-a)} \equiv a^a(-a)^{(p-a)} \equiv (-1)^{(p-a)}a^p \equiv (-1)^{(p-a)}a \pmod{p},$$

it follows that

$$a^a(p-a)^{(p-a)} \equiv \pm\, a \pmod{p}, \qquad \begin{array}{l} +\text{if } a \text{ is odd} \\ -\text{if } a \text{ is even.} \end{array}$$

Hence,

$$\prod_{a=1}^{(p-1)/2} a^a(p-a)^{(p-a)} \equiv (-1)^q\left(\frac{p-1}{2}\right)! \pmod{p},$$

where q is the number of positive even integers $\leqq (p-1)/2$;

$$q=\frac{p-1}{4} \quad \text{if } p=4x+1, \qquad \text{and} \qquad q=\frac{p-3}{4} \quad \text{if } p=4x+3.$$

Remembering that

$$(-1)^{(p-1)/2}\left(\frac{p-1}{2}\right)^2! \equiv (p-1)! \equiv -1 \pmod{p},$$

one gets

THEOREM II.

$$\left\{(p-1)^{(p-1)}(p-2)^{(p-2)} \cdots 2^2 1^1\right\}^2 \equiv -1 \pmod{p} \quad \text{if } p=4x+1.$$

$$(p-1)^{(p-1)}(p-2)^{(p-2)} \cdots 2^2 1^1 \equiv \pm 1 \pmod{p} \quad \text{if } p=4x+3,$$

or, more exactly, $\equiv(-1)^{[(p-3)/4]+\nu} \pmod{p}$, *where* ν *= the number of quadratic nonresidues* $<\frac{1}{2}p$.

Thus one has an analogue of Wilson's Theorem for n^n.

Another result may be obtained as follows. If $n^n \equiv a \pmod{p}$, then $x^n \equiv a \pmod{p}$ has a solution, which implies that

$$a^{(p-1)/(n,p-1)} \equiv 1 \pmod{p}.$$

Now if a is a primitive root of p then $(n, p-1)=1$. Hence, if a is a primitive root of p, it can only be a residue of n^n with n such that $(n, p-1)=1$.

One final result. If p is reducible for integers $c, b<[\sqrt{p-1}]$, c and b distinct or not, then the following method can be used to find another integer for which p is reducible. Since

$$c^c \equiv 1 \pmod{p} \quad \text{and} \quad b^b \equiv 1 \pmod{p}, \quad 1<b, \quad c<[\sqrt{p-1}],$$

then since

$$(cb)^{cb} \equiv (c^c)^b(b^b)^c \equiv 1^b 1^c \equiv 1 \pmod{p},$$

it follows that $(cb)^{cb} \equiv 1 \pmod{p}$, where $cb<p-1$.

ON FAMILIES OF TOPOLOGIES FOR A SET

NORMAN LEVINE, Ohio State University

1. Throughout this paper, X and Δ will denote nonempty sets, Δ consisting of two or more elements. For each $\alpha \in \Delta$, let $\mathfrak{J}_\alpha$ be a topology for X. The family $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ generates a topology $\mathfrak{J}$ for X in the following natural way: a subset O of X is in $\mathfrak{J}$ iff for each $x \in O$, there exist $\alpha_1, \cdots, \alpha_n$ in Δ and $U_i \in \mathfrak{J}_{\alpha_i}$ such that $x \in U_1 \cap \cdots \cap U_n \subset O$. It is easy to verify that (i) $\mathfrak{J}$ is a topology for X, (ii) $\mathfrak{J}_\alpha \subset \mathfrak{J}$ for each $\alpha \in \Delta$ and (iii) $\mathfrak{J}$ is minimal relative to properties (i) and (ii).

2. We will be concerned with this general question: If for each $\alpha \in \Delta$, $(X, \mathfrak{J}_\alpha)$ has some property P, under what circumstances will $(X, \mathfrak{J})$ also have property P?

3. For each $\alpha \in \Delta$, let $X_\alpha = X$ and let $(X^*, \mathfrak{J}^*) = X\{(X_\alpha, \mathfrak{J}_\alpha) : \alpha \in \Delta\}$. Let D^* be the diagonal in X^*, that is, $x^* \in D^*$ iff $x^* : \Delta \to X$ is a constant. For each $\alpha \in \Delta$, P_α will denote the projection from X^* onto X_α.

4. THE FUNDAMENTAL THEOREM. *$(X, \mathfrak{J})$ is homeomorphic to $(D^*, \mathfrak{J}^* \cap D^*)$ (see sections* 1, 3).

Proof. Let $h: D^* \to X$ as follows: $h(x^*) = x^*(\alpha)$ for all $\alpha \in \Delta$; h is clearly single-valued, one-to-one and onto. We show first that h is continuous. Suppose $y^* \in D^*$ and that $h(y^*) \in O \in \mathfrak{J}$. Then there exist $\alpha_1, \cdots, \alpha_n$ in Δ and $U_i \in \mathfrak{J}_{\alpha_i}$ for which $h(y^*) \in U_1 \cap \cdots \cap U_n \subset O$ (see section 1). Then $P_{\alpha_i}(y^*) = y^*(\alpha_i) = h(y^*) \in U_i$ for $i = 1, \cdots, n$. Thus $y^* \in P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap P_{\alpha_n}^{-1}[U_n] \cap D^* \in \mathfrak{J}^* \cap D^*$. But

$$\begin{aligned}
&h\{P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap P_{\alpha_n}^{-1}[U_n] \cap D^*\} \\
&\quad = h\{D^* \cap P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap D^* \cap P_{\alpha_n}^{-1}[U_n]\} \\
&\quad = P_{\alpha_1}\{D^* \cap P_{\alpha_1}^{-1}[U_1]\} \cap \cdots \cap P_{\alpha_n}\{D^* \cap P_{\alpha_n}^{-1}[U_n]\} \\
&\quad \subset P_{\alpha_1}P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap P_{\alpha_n}P_{\alpha_n}^{-1}[U_n] \\
&\quad = U_1 \cap \cdots \cap U_n \subset O.
\end{aligned}$$

Thus h is continuous.

Next we show that h is an open map. Let $x \in h(O^* \cap D^*)$ where $O^* \in \mathfrak{J}^*$. Then $x = h(x^*)$ for some $x^* \in O^* \cap D^*$. There exist then $\alpha_1, \cdots, \alpha_n$ in Δ and $U_i \in \mathfrak{J}_{\alpha_i}$ for which

$$x^* \in P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap P_{\alpha_n}^{-1}[U_n] \cap D^* \subset O^* \cap D^*$$

(see section 3). But $x = h(x^*) = x^*(\alpha_i) = P_{\alpha_i}(x^*) \in U_i$ for each i. Thus $x \in U_1 \cap \cdots \cap U_n \in \mathfrak{J}$. The proof will be complete when we show that $U_1 \cap \cdots \cap U_n \subset h(O^* \cap D^*)$. Now

$$\begin{aligned}
h^{-1}[U_1 \cap \cdots \cap U_n] &= h^{-1}[U_1] \cap \cdots \cap h^{-1}[U_n] \\
&= D^* \cap P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap D^* \cap P_{\alpha_n}^{-1}[U_n] \subset O^* \cap D^*.
\end{aligned}$$

Since h is one-to-one and onto, $U_1 \cap \cdots \cap U_n \subset h(O^* \cap D^*)$. Thus h is a homeomorphism.

5. COROLLARY. *For each* $\alpha \in \Delta$, $(X, \mathfrak{I}_\alpha)$ *is a continuous, one-to-one image of* $(D^*, \mathfrak{I}^* \cap D^*)$ (*see section* 3).

Proof. The map $h: D^* \to X$ in section 4 is a homeomorphism (relative to $\mathfrak{I}^* \cap D^*$ and $\mathfrak{I}$). Since $\mathfrak{I}_\alpha \subset \mathfrak{I}$ for each $\alpha \in \Delta$, h is continuous relative to $\mathfrak{I}^* \cap D^*$ and $\mathfrak{I}_\alpha$.

6. COROLLARY. *Let P be a property which is invariant under continuous, one-to-one, onto maps, e.g., connectness, compactness dense-in-itself, separability, etc. If* $(X, \mathfrak{I})$ *has property P, then* $(X, \mathfrak{I}_\alpha)$ *also has property P for each* $\alpha \in \Delta$.

7. A property P of a space is termed hereditary (closed hereditary) iff every subspace (closed subspace) of a space with property P also has property P; P is called productive (countably productive) iff the product (countable product) of spaces enjoying P also has property P.

8. COROLLARY. *Suppose P is a property which is both productive and hereditary* (*section* 7), *e.g.*, T_1, T_2, *regular, completely regular, Tychonoff, etc. If each* $(X, \mathfrak{I}_\alpha)$ *has property P, then* $(X, \mathfrak{I})$ *also has property P.*

Proof. $(D^*, \mathfrak{I}^* \cap D^*)$ has property P since it is a subspace in the product space. Hence $(X, \mathfrak{I})$ has property P by section 4.

9. COROLLARY. *Let P be a property which is hereditary and countably productive* (*see section* 7), *e.g., metrizability, first axiom, second axiom, etc. If Δ is countable and* $(X, \mathfrak{I}_\alpha)$ *has property P for each* $\alpha \in \Delta$, *then* $(X, \mathfrak{I})$ *has property P.*

10. COROLLARY. *Let P be a property which is productive and closed-hereditary* (*see section* 7), *e.g., compactness. If* $(X, \mathfrak{I}_\alpha)$ *has property P for each* $\alpha \in \Delta$ *and if* D^* *is closed, then* $(X, \mathfrak{I})$ *has property P* (*see sections* 1, 3).

11. THEOREM. *If* D^* *is closed, then* $(X, \mathfrak{I})$ *is* T_2 (*see sections* 1, 3).

Proof. Let $x \neq y$. Take $\alpha^* \in \Delta$ and let $x^*: \Delta \to X$ as follows: $x^*(\alpha^*) = x$ and $x^*(\alpha) = y$ for all $\alpha \neq \alpha^*$ (recall that Δ consists of two or more elements). Then $x^* \notin D^*$ and, since we presume that D^* is closed, there exist $\alpha_1, \cdots, \alpha_n$ in Δ and $U_i \in \mathfrak{I}_{\alpha_i}$ for which $x^* \in P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap P_{\alpha_n}^{-1}[U_n] \subset \mathcal{C}^* D^*$, $\mathcal{C}^*$ denoting the complement operator in X^*. We remark, first that $n \geqq 2$. If $n = 1$, take $q \in U_1$ and let $z^*: \Delta \to X$ as follows: $z^*(\alpha) = q$ for all $\alpha \in \Delta$. Then $z^* \in P_{\alpha_1}^{-1}[U_1]$, but $z^* \notin \mathcal{C}^* D^*$ which is a contradiction. Case 1. $\alpha^* \neq \alpha_i$ for each i. Then $y = x^*(\alpha_i) = P_{\alpha_i}(x^*) \in U_i$. Let $y^*: \Delta \to X$ as follows: $y^*(\alpha) = y$ all $\alpha \in \Delta$. Then $y^* \in P_{\alpha_1}^{-1}[U_1] \cap \cdots \cap P_{\alpha_n}^{-1}[U_n]$, but $y^* \notin \mathcal{C}^* D^*$ which is a contradiction. Case 2. α^* is one of the α_i, say α_1. Then

$$x^* \in P_{\alpha_1}^{-1}[U_1] \cap P_{\alpha_2}^{-1}[U_2] \cap \cdots \cap P_{\alpha_n}^{-1}[U_n] \subset \mathcal{C}^* D^*.$$

But $x = x^*(\alpha^*) \in U_1$ and $y = x^*(\alpha_i) \in U_i$ for $i = 2, \cdots, n$. Then $x \in U_1 \in \mathfrak{I}$ and

$y \in U_2 \cap \cdots \cap U_n \in \mathfrak{J}$. It is left to the reader to show that U_1 and $U_2 \cap \cdots \cap U_n$ are disjoint.

12. The converse of the theorem in section 11 is false (see example 17.3). A sufficient condition for D^* to be closed in $(X^*, \mathfrak{J}^*)$ is given in section 14.

13. Let $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ be a family of topologies for X. Then $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ has the *chain property* iff for $\alpha, \beta \in \Delta$, then $\mathfrak{J}_\alpha \subset \mathfrak{J}_\beta$ or $\mathfrak{J}_\beta \subset \mathfrak{J}_\alpha$.

14. THEOREM. *Let $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ be a chain of T_2-topologies for X. Then D^* is closed in* $(X^*, \mathfrak{J}^*)$ (*see sections* 1, 3, 13).

Proof. Let $x^* \notin D^*$. Then there exist $\alpha, \beta \in \Delta$ such that $x^*(\alpha) \neq x^*(\beta)$. Suppose that $\mathfrak{J}_\alpha \subset \mathfrak{J}_\beta$. Now there exist disjoint U and V in $\mathfrak{J}_\alpha$ for which $x^*(\alpha) \in U$ and $x^*(\beta) \in V$. Since $V \in \mathfrak{J}_\beta$ and since $x^* \in P_\alpha^{-1}[U] \cap P_\beta^{-1}[V] \subset \mathfrak{C}^* D^*$, it follows that D^* is closed.

15. THEOREM. *Let $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ be a chain of topologies for X and let $O \subset X$. Then $O \in \mathfrak{J}$ iff for each $x \in O$, there exists a $U \in \mathfrak{J}_\alpha$ for some $\alpha \in \Delta$ such that $x \in U \subset O$,* (*see sections* 1, 13).

Proof. The sufficiency follows from the definition of $\mathfrak{J}$. To show the necessity, let $x \in O \in \mathfrak{J}$. Then there exist $\alpha_1, \cdots, \alpha_n$ in Δ and $U_i \in \mathfrak{J}_{\alpha_i}$ such that $x \in U_1 \cap \cdots \cap U_n \subset O$. Suppose that the notation is chosen so that $\mathfrak{J}_{\alpha_1} \subset \mathfrak{J}_{\alpha_2} \subset \cdots \subset \mathfrak{J}_{\alpha_n}$. Let $U = U_1 \cap \cdots \cap U_n$. U clearly is in $\mathfrak{J}_{\alpha_n}$.

16. THEOREM. *Let $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ be a chain of connected topologies for X. If $(X, \mathfrak{J})$ is compact, then $(X, \mathfrak{J})$ is connected* (*see example* 17.3 *in this connection*).

Proof. Suppose that $(X, \mathfrak{J})$ is not connected. Then there exist O_1 and O_2 disjoint and nonempty in $\mathfrak{J}$ for which $X = O_1 \cup O_2$. For each $x \in O_1$, there exists a $U_x \in \bigcup\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$ such that $x \in U_x \subset O_1$ (see section 15). Since $O_1 = \bigcup\{U_x : x \in O_1\}$ and since O_1 is closed and therefore compact in $(X, \mathfrak{J})$, it follows that $O_1 = U_{x_1} \cup \cdots \cup U_{x_n}$ and hence $O_1 \in \mathfrak{J}_\alpha$ for some $\alpha \in \Delta$. Similarly $O_2 \in \mathfrak{J}_\beta$ for some $\beta \in \Delta$. Suppose now that $\mathfrak{J}_\beta \subset \mathfrak{J}_\alpha$. Then $(X, \mathfrak{J}_\alpha)$ is disconnected which is a contradiction.

17. In this last section we give some examples to illustrate the possible limitation of properties which $\mathfrak{J}$ may inherit from $\{\mathfrak{J}_\alpha : \alpha \in \Delta\}$. In fact, in each of the following examples, will consist of exactly two objects, 1 and 2.

Example 17-1. Let $X = [0, 1]$ and let $\mathfrak{J}_1$ consist of the empty set and all subsets of $(0, 1]$ together with all sets containing 0 which have finite complements. Let $\mathfrak{J}_2$ consist of the empty-set, all subsets of $[0, 1)$ and all sets containing 1 which have finite complements. Then $(X, \mathfrak{J}_1)$ and $(X, \mathfrak{J}_2)$ are each compact, countably compact and Lindelöf. But $(X, \mathfrak{J})$ is discrete and thus possesses none of these properties.

Example 17-2. Let X be the reals and let $\mathfrak{J}_1$ be the usual topology. Let $\mathfrak{J}_2$ consist of the empty-set, X and all subsets of the rationals. Now $(X, \mathfrak{J}_1)$ is locally compact and $(X, \mathfrak{J}_2)$ is compact and thus also locally compact. But

$(X, \mathfrak{J})$ is not locally compact. To see this, let $\sqrt{2} \in O \in \mathfrak{J}$ and suppose $c(O)$ is compact in $(X, \mathfrak{J})$, c denoting the closure operator. Then there exist rationals $r<s$ such that $\sqrt{2} \in (r, s) \subset O \subset c(O)$. But (r, s) is both open and closed in $(X, \mathfrak{J})$ and thus is compact since $c(O)$ is presumed to be compact. But (r, s) is not compact in $(X, \mathfrak{J}_1)$ and thus cannot be compact in $(X, \mathfrak{J})$ since $\mathfrak{J}_1 \subset \mathfrak{J}$.

Example 17-3. Let $X: a, b, \mathfrak{J}_1: \varnothing, (a), X$ and $\mathfrak{J}_2: \varnothing, (b), X$. Then $(X, \mathfrak{J}_1)$ and $(X, \mathfrak{J}_2)$ are connected, but $(X, \mathfrak{J})$ is discrete and thus disconnected. Note that $\{\mathfrak{J}_i: i=1, 2\}$ is not a chain (see section 16).

Example 17-4. Let $X=\{1, 2, \cdots, n, \cdots\}$ and let $\mathfrak{J}_1$ consist of the empty-set, X, and all subsets of $\{2, 3, \cdots, n, \cdots\}$. Let $\mathfrak{J}_2$ consist of the empty-set together with all subsets of X whose complements are finite. It is left to the reader to verify that each $(X, \mathfrak{J}_i)$ is locally connected, but that $(X, \mathfrak{J})$ is not locally connected at 1.

Example 17-5. Let $X: a, b, c, \mathfrak{J}_1: \varnothing, (a, c), X$ and $\mathfrak{J}_2: \varnothing, (a, b), X$. It is easy to verify that each $(X, \mathfrak{J}_i)$ is normal, but that $(X, \mathfrak{J})$ is not.

Example 17-6. Let X be the reals, and $\mathfrak{J}_1$ the topology generated by right semi-open intervals and $\mathfrak{J}_2$ the topology generated by left semi-open intervals. Each $(X, \mathfrak{J}_i)$ is separable, but $(X, \mathfrak{J})$ is discrete and hence not separable.

Reference

1. John L. Kelley, General Topology, Van Nostrand, Princeton, N. J., 1955.

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

A CHARACTERIZATION AND GENERALIZATION OF SEMI-METRIZABILITY

D. E. SANDERSON, Iowa State University and B. T. SIMS, San Jose State College

1. Introduction. A semi-metric ρ on a set S is a nonnegative, real-valued function, defined on $S \times S$ and such that (i) $\rho(x, y)=0$ if and only if $x=y$; (ii) $\rho(x, y)=\rho(y, x)$ for all $x, y \in S$. A topological space $\langle S, \tau\rangle$ is semi-metrizable if and only if there exists a semi-metric ρ on S such that for each $p \in S$ the collection $\{S_\rho(p; r) \mid r>0\}$ is a local base at p for τ, where $S_\rho(p; r)=\{x \in S \mid \rho(p, x) < r\}$. Such a semi-metric is said to be admissible on $\langle S, \tau\rangle$. Observe that ρ is a metric on S and metrizes $\langle S, \tau\rangle$ if, in addition, $\rho(x, y) \leqq \rho(x, z)+\rho(z, y)$ for all $x, y, z \in S$. It is well known that every metrizable space is a completely normal T-space, but that semi-metrizable spaces are T_1-spaces which may not even be Hausdorff. A good discussion of the basic properties of semi-metric spaces is contained in [1].

In this paper, we present a characterization of semi-metrizable spaces in terms of "indexed" neighborhoods. It is similar to those given by Davis in [2] for topological spaces and by Hall and Spencer in [3] for metrizable spaces. We

generalize the concept of semi-metrizability by introducing the notion of an a-metrizable space which requires no T_i-separation. An example of an a-metrizable space which is not semi-metrizable is given in section 3, and Theorem 2 asserts that the concepts of a-metrizability and semi-metrizability are equivalent for the class of T_0-spaces.

2. Semi-metrizable spaces. For our semi-metrization theorem, we shall need the following lemma on indexed neighborhood systems for a topological space $\langle S, \tau\rangle$. Let $\{N_i \mid i\in I^+\}$ be a family of functions indexed by the positive integers, I^+, which assign to each $x\in S$ a subset $N_i(x)$ of S and consider the following properties:

(i) for each $x\in S$, the collection $\{N_i(x)\mid i\in I^+\}$ is a local base at x;

(ii) for each $i\in I^+$, $x\in N_i(y)$ implies $y\in N_i(x)$;

(iii) for $m, n\in I^+$, there exists a $k\in I^+$ such that for all $x\in S$, $N_k(x)\subset N_m(x)\cap N_n(x)$;

(iii)$'$ for each $x\in S$ and $m, n\in I^+$, $N_m(x)\subset N_n(x)$ if $m>n$.

LEMMA. *Condition* (iii)$'$ *implies* (iii) *and if there exists a family* $\{N_i\mid i\in I^+\}$ *satisfying* (i)–(iii), *then there is a family* $\{N_i'\mid i\in I^+\}$ *satisfying* (i), (ii) *and* (iii)$'$.

Proof. Choosing k to be the larger of m and n suffices for the first part. The second part is proved by setting $N_1'=N_1$ and $N_{i+1}'=N_k$ (for $i\in I^+$) where k is the first positive integer such that $N_k(x)\subset N_i'(x)\cap N_{i+1}(x)$ for all x.

THEOREM 1. *A necessary and sufficient condition that a T_1-space $\langle S, \tau\rangle$ be semi-metrizable is that there exists a family $\{N_i\mid i\in I^+\}$ of functions which assign to each $x\in S$ a subset $N_i(x)$ of S such that conditions* (i)–(iii) *of this section are satisfied.*

Proof. Necessity. Let ρ be any admissible semi-metric on $\langle S, \tau\rangle$, and let $N_i(x)=S_\rho(x; 1/i)$ for each $i\in I^+$ and each $x\in S$. It is easily seen that $\{N_i\mid i\in I^+\}$ satisfies (i), (ii) and (iii)$'$. By the Lemma, it satisfies (iii) also.

Sufficiency. Let $\{N_i\mid i\in I^+\}$ be any family of functions satisfying (i)–(iii). By the Lemma, there exists a family $\{N_i'\mid i\in I^+\}$ of functions satisfying (i), (ii) and (iii)$'$. Then

$$\bigcap_{i\in I^+} N_i'(x) = \{x\}.$$

Define a semi-metric ρ on S as follows:

(i) $\rho(x, y)=0$ if $x=y$; (ii) $\rho(x, y)=\sup\{1/i\mid y\notin N_i'(x),\ i\in I^+\}$ if $x\neq y$.

For each $x\in S$ and each $i\in I^+$, $S_\rho(x; 1/i)=N_i'(x)$, and thus $\langle S, \tau\rangle$ is semi-metrizable since $\{N_i'(x)\mid i\in I^+\}$ is a local base at x.

It is possible that the existence of a semi-metric with a countable distance set, such as that which we have just constructed, could be useful in the construction of induction proofs in a semi-metrizable space. With the introduction of such a semi-metric on a non-totally disconnected space, however, we lose all manner of continuity of the semi-metric.

3. *A*-metrizable spaces. In [**3**, p. 65] Hall and Spencer give the following example of a Hausdorff space which is not regular (hence not metrizable), but is locally metrizable. Let S be the closed upper half-plane in E^2, and let L denote the x-axis. For each $p\in S$ and each $r>0$, define a neighborhood $N_r(p)$ as follows:

(i) $N_r(p)=S_d(p;r)\cap S$ if $p\in S-L$;

(ii) $N_r(p)=[S_d(p;r)\cap(S-L)]\cup\{p\}$ if $p\in L$,

where d is the usual metric for E^2 and $S_d(p;r)=\{x\in E^2 \mid d(p,x)<r\}$. It is easily seen that a topology τ is generated for S by the collection $\{N_r(p)\mid p\in S, r>0\}$, and that $\langle S,\tau\rangle$ has the following properties: (i) it satisfies the first axiom of countability, but not the second; (ii) it is separable, but not hereditarily separable. Define a metric ρ on S as follows:

(i) $\rho(p,q)=0$ if and only if $p=q$;

(ii) $\rho(p,q)=d(p,q)+\frac{1}{2}$ if $p\notin L$ or $q\notin L$;

(iii) $\rho(p,q)=d(p,q)+1$ if $p, q\in L$.

For each $p\in S$, it is easily seen that $S_\rho(p;r)=N_{r-1/2}(p)$, $\frac{1}{2}<r\leqq 1$. Thus, although ρ does not metrize $\langle S,\tau\rangle$, we observe that for each $p\in S$ the collection $\{S_\rho(p;r)\mid r>\frac{1}{2}\}$ is a local base at p for τ. Abstracting from this example, we are led to a natural formulation of the concept of "a-metrizability."

DEFINITION. *Given a real number $a\geqq 0$, a topological space $\langle S,\tau\rangle$ is a-metrizable if and only if there exists a metric ρ on S such that for each $p\in S$, the collection $\{S_\rho(p;r)\mid r>a\}$ is a local base at p for τ. ρ is called an a-metric for $\langle S,\tau\rangle$.*

Since an a-metrizable space is "almost" metrizable, the following properties either follow immediately from the definition or may be established by imitating the proofs of their metric analogues (e.g., see [**3**]).

I. $\langle S,\tau\rangle$ is metrizable if and only if it is 0-metrizable.

II. A-metrizability is a hereditary property.

III. If $\langle S,\tau\rangle$ is an a-metrizable T_0-space, then it is a T_1-space.

IV. If $\langle S,\tau\rangle$ is a-metrizable, it satisfies the first axiom of countability.

V. Every a-metrizable space has a bounded a-metric.

VI. A-metrizability is a topological property.

In the exposition which follows, we consider the relationship between the concepts of "a-metrizability" and "semi-metrizability." Our next theorem and example establish that a-metrizability is a generalization of semi-metrizability for topological spaces possessing no T_i-separation.

THEOREM 2. *A necessary and sufficient condition that a topological space $\langle S,\tau\rangle$ be semi-metrizable is that it be an a-metrizable T_0-space.*

Proof. Suppose ρ is an a-metric for $\langle S,\tau\rangle$. For $x, y\in S$, let (i) $d(x,y)=0$ if and only if $x=y$ and (ii) $d(x,y)=\rho(x,y)-a$ if $x\neq y$ ($\rho(x,y)>a$ if $\langle S,\tau\rangle$ is T_0 and hence $d(x,y)>0$). Since $S_d(p;r-a)=S_\rho(p;r)$ for $r>a$ and each $p\in S$, d is an admissible semi-metric on $\langle S,\tau\rangle$.

Conversely, assume that d is a bounded, admissible semi-metric on $\langle S, \tau\rangle$, and set

$$a = \sup\{d(x, y) - d(x, z) - d(z, y) \mid x, y, z \in S\}.$$

For $x, y \in S$, let (i) $\rho(x, y) = 0$ if and only if $x = y$ and (ii) $\rho(x, y) = d(x, y) + a$ if $x \neq y$. It is easily shown that ρ is a metric on S and, since $S_\rho(p; r+a) = S_d(p; r)$ for $r > 0$ and each $p \in S$, ρ is an a-metric for $\langle S, \tau\rangle$.

Example. To verify that a-metrizability is a generalization of semi-metrizability, we give an example of a topological space which is a-metrizable but not semi-metrizable. Let $S = \{b, c, d\}$ and $\tau = \{\varnothing, S, \{c, d\}, \{b\}\}$. $\langle S, \tau\rangle$ is not a T_0-space (hence not semi-metrizable). Define a metric ρ on S as follows:

(i) $\rho(b, c) = \rho(c, b) = \rho(b, d) = \rho(d, b) = 2$;

(ii) $\rho(c, d) = \rho(d, c) = 1$.

ρ is an a-metric ($a = 1$) for $\langle S, \tau\rangle$, since if $1 < r \leqq 2$, then $S_\rho(b; r) = \{b\}$ and $S_\rho(c; r) = S_\rho(d; r) = \{c, d\}$.

Note that since there are T_0-spaces which satisfy the first axiom of countability but are not T_1-spaces, properties III, IV, Theorem 2 and the above example imply that the class of a-metrizable spaces occupies a position strictly intermediate to the classes of semi-metrizable and first countable spaces.

REMARK. In constructing an a-metric ρ from the semi-metric d in the proof of Theorem 2, we set

$$a = \sup\{d(x, y) - d(x, z) - d(z, y) \mid x, y, z \in S\},$$

but this choice of "a" is not unique. In fact, if ρ is an admissible a-metric for $\langle S, \tau\rangle$ and $b > 0$, one can set $\rho'(x, y) = \rho(x, y) + b$ if $x \neq y$ ($\rho'(x, x) = 0$ for all $x \in S$) and ρ' will be an admissible $(a+b)$-metric for $\langle S, \tau\rangle$. That is, although ρ' clearly induces the discrete topology on S, the ρ'-spheres of radius exceeding $a+b$ form a basis for τ. At the other end of the spectrum the number a cannot be reduced below the value used in the proof of Theorem 3, but of course one can diminish the "scale," e.g. by defining

$$\rho_k(x, y) = \frac{\rho'(x, y)}{k(\rho'(x, y) + 1)}$$

for any $k \in I^+$. This gives an admissible a_k-metric for $\langle S, \tau\rangle$ with

$$a_k = \frac{a + b}{k(a + b + 1)}$$

as small as one may wish, and the associated semi-metric has correspondingly small deficiency in the triangle inequality. This is no surprise, however, since all distances are then small. Observe that the (pointwise) limit function $\rho = \lim_{k\to\infty} \rho_k$ is a pseudometric inducing the indiscrete topology of S.

References

1. Seminar on semi-metric spaces, Summary of lectures and seminars, Summer Institute on Set-theoretic Topology, Amer. Math. Soc., Madison (1955) 58–64.
2. A. S. Davis, Indexed systems of neighborhoods for general topological spaces, this MONTHLY, 68 (1961) 886–893.
3. D. W. Hall and G. L. Spencer, Elementary Topology, Wiley, New York, 1955.

QUADRIC SURFACES WHICH TOUCH THE EDGES OF A TETRAHEDRON

O. BOTTEMA, Delft, Holland

1. Introduction. We consider the set S of ∞^3 quadrics Q which touch the six edges of a given tetrahedron $A_1A_2A_3A_4$ and study some of its *affine* properties. The major result deals with the centers of Q which are shown not to fill up the whole space but to cover a region bordered by the rational quartic which is known as Steiner's *Roman surface* (see [1], p. 221, or [2], Chapter 7).

2. The equation of the set. We take A_i as the frame of reference for a system of homogeneous barycentric coordinates x_i $(i=1, 2, 3, 4)$; the plane V at infinity has the equation $\sum x_i=0$. A quadric $\sum a_{ij}x_ix_j=0$ is tangent to the edge A_kA_l if $a_{kl}^2=a_{kk}a_{ll}$ and therefore a quadric Q has an equation

$$\sum_i u_i^2x_i^2 + 2\sum_{i>j} \pm u_iu_jx_ix_j = 0 \tag{1}$$

with u_i as parameters. As the system is unchanged if we replace a parameter u_k by $-u_k$ we may suppose without loss of generality that the signs of u_1u_2, u_1u_3, u_1u_4 are negative. Q is a double plane if the three remaining signs are all positive and a cylinder if either one or two of them are positive. If we exclude these special cases the only one left is

$$\sum u_i^2x_i^2 - 2\sum u_iu_jx_ix_j = 0 \tag{2}$$

the discriminant of which reads $\Delta=-16u_1^2u_2^2u_3^2u_4^2$. Therefore: *the set S of non-singular quadrics touching the edges of $A_1A_2A_3A_4$ is given by* (2), *in which all the parameters u_i are different from zero.*

3. Configuration of the points of contact. We introduce the variables $v_i=u_i^{-1}$ $(i=1, 2, 3, 4)$. Then the point of contact P_{12} on A_1A_2 is $(v_1, v_2, 0, 0)$, on A_1A_3 it is $P_{13}=(v_1, 0, v_3, 0)$, and so on. Hence *the lines $P_{12}P_{34}$, $P_{13}P_{42}$ and $P_{14}P_{23}$ joining the points of contact on opposite edges pass through the point $P(v_1, v_2, v_3, v_4)$* which will be called the *representative* of Q. Conversely, each point not on the surface of the tetrahedron is the representative of *one* quadric of the set.

4. Classification of the quadrics. As $\Delta<0$, no ruled quadrics belong to S. Hence the only affine types are: the *ellipsoid*, the *elliptic paraboloid* and the *hyperboloid of two sheets*. In order to distinguish between them we determine

the *center* $C(y_1, y_2, y_3, y_4)$ of Q, that is, the pole of the plane V. The polar plane of C has the equation

$$\sum u_1 x_1(u_1y_1 - u_2y_2 - u_3y_3 - u_4y_4) = 0, \tag{3}$$

and it coincides with $\sum x_i = 0$ if $u_1(u_1y_1 - u_2y_2 - u_3y_3 - u_4y_4) = p$ and three similar equations for the same constant p are satisfied. From this it follows that $-2\sum u_iy_i = p\sum v_i$ and so we find

$$y_i = v_i(v - v_i), \tag{4}$$

where $2v = v_1+v_2+v_3+v_4$. Hence *the coordinates of the center C are quadratic functions of the coordinates of the representative point P.* The center is at infinity if $\sum y_i = 0$ or

$$F(v_1, v_2, v_3, v_4) \equiv \sum_i v_i^2 - 2\sum_{i>j} v_iv_j = 0. \tag{5}$$

This is the equation of the quadric (evidently an ellipsoid) which touches the edges at their midpoints and which we shall denote by F. Therefore, *Q is a paraboloid if its representative $P(Q)$ is on F* and we verify easily that *Q is an ellipsoid if its representative is inside F and a hyperboloid if it is outside F.*

5. Quadrics with a given center. In order to have a given point as its center a quadric has to satisfy three conditions and so the question arises whether there exists a quadric of S with prescribed center C. This means, analytically, whether we are able to solve the equations (4) if the y_i are given and the v_i are unknowns.

From (4) it follows that $y_i - y_j = (v_i - v_j)(v - v_i - v_j)$. Hence we have either

$$v_1 + v_2 - v_3 - v_4 = 0 \tag{6}$$

or

$$(y_3 - y_4)(v_1 - v_2) + (y_1 - y_2)(v_3 - v_4) = 0, \tag{7}$$

which are both *linear* equations. From (4) and (6) we deduce that $y_1 = y_2 = v_1v_2$, $y_3 = y_4 = v_3v_4$ and these satisfy (7), so that we may restrict ourselves to the latter. There are two similar equations:

$$(y_4 - y_2)(v_1 - v_3) + (y_1 - y_3)(v_4 - v_2) = 0, \tag{8}$$

$$(y_2 - y_3)(v_1 - v_4) + (y_1 - y_4)(v_2 - v_3) = 0. \tag{9}$$

The rank of the system (7), (8) and (9) is *two*; they have an infinity of solutions. One of them is $v_1 = v_2 = v_3 = v_4$ and another is

$$v_i = y_i(y - y_i), \tag{10}$$

where $2y$ stands for $y_1+y_2+y_3+y_4$. The solutions are therefore

$$v_i = y_i(y - y_i) + \lambda, \tag{11}$$

where λ is a parameter. The system (4) is equivalent to three quadratic equations. We take one of them, for instance

$$y_2v_1(v - v_1) - y_1v_2(v - v_2) = 0. \tag{12}$$

If we eliminate v_i from (11) and (12) and use the notation of (5), we obtain a quadratic equation for λ:

$$4\lambda^2 - \lambda F(y_1, y_2, y_3, y_4) + 4y_1y_2y_3y_4 = 0. \tag{13}$$

Thus

$$\lambda = \tfrac{1}{8}(F \pm \sqrt{D}), \tag{14}$$

where

$$D = F^2(y_1, y_2, y_3, y_4) - 64y_1y_2y_3y_4. \tag{15}$$

Hence there are *two* quadrics which touch the edges of a given tetrahedron and which have a prescribed center $C(y_1, y_2, y_3, y_4)$. The coordinates of their representatives are given by (11), (14) and (15) as functions of y_i. The quadrics are real and distinct if $D>0$; they coincide if $D=0$; they have imaginary equations if $D<0$. The two representatives are on a line through the centroid (1, 1, 1, 1).

Real quadrics of the set S have their centers in that region of the space whose points satisfy $D\geqq 0$. The bordering surface is $D=0$. This, being equivalent to

$$\sqrt{y_1} \pm \sqrt{y_2} \pm \sqrt{y_3} \pm \sqrt{y_4} = 0,$$

is the standard equation of the rational quartic surface ("Roman surface") which was discovered by Jacob Steiner about 1840 and has many other remarkable properties.

References

1. H. F. Baker, Principles of geometry, vol. 3, Cambridge University Press, 1923.
2. C. M. Jessop, Quartic surfaces with singular points, Cambridge University Press, 1916.

ON THE BOUNDEDNESS AND STABILITY OF SOLUTIONS OF DIFFERENTIAL-DIFFERENCE EQUATIONS WITH LAG

YAU-CHUEN WONG, Chu Hai College, Hong Kong

In this paper, we discuss the boundedness of all solutions of the following differential-difference equation

$$\frac{du(t)}{dt} = b_1(t)u(t) + b_2(t)u(t-\lambda) + F(t), \tag{1}$$

where $b_i(t)$ $(i=1, 2)$ and $F(t)$ are real functions of a real variable t, and $\lambda>0$ is called the *lag*.

First, we investigate that the trivial solution which is asymptotically stable with respect to the linear homogeneous equation

$$\frac{du(t)}{dt} = b_1(t)u(t) + b_2(t)u(t-\lambda). \tag{1'}$$

We can then establish the boundedness of all solutions of the equation (1). Chian Hod-Ming [1] employed the method of Liapunoff's V-function to investigate the case in which $b_i(t)$ $(i=1, 2)$ are constants in the equations (1') and (1).

For the definition of the stability of differential-difference equations, we refer to the treatise of Aliskolish [2, 3], or Krasovskii [4], or Myshkis [5].

The initial set E_{t_0} of the equation (1) (or (1')) is the interval $t_0-\lambda \leqq t \leqq t_0$. On E_{t_0} we specify the initial function of (1) (or (1')), i.e.,

$$u(t)\,\big|_{t_0-\lambda\leqq t\leqq t_0} = g(t),$$

where $g(t)$ is a known bounded function.

THEOREM 1. *If the trivial solution of the linear first approximation of* (1)' *is stable,* (*namely the integral* $\int_{t_0}^t b_1(x)dx$ *is bounded, for all* $t \geqq t_0$) *and*

(i) $|u(t)| \leqq 1$ *when* $t \geqq t_0$,

(ii) $\int_{t_0}^t |b_2(x+\lambda)|\,dx$ *is bounded, when* $t \geqq t_0$

then the trivial solution of (1)' *is asymptotically stable.*

Proof. First, the equation

$$\frac{du(t)}{dt} = b_1(t)u(t)$$

has the solution

$$u(t) = c\cdot\exp\left(\int_{t_0}^t b_1(x)dx\right).$$

From this, we obtain the solution of (1)'

$$u(t) = \left[\int_{t_0}^t b_2(x)u(x)u(x-\lambda)\cdot\exp\left(-\int_{t_0}^x b_1(y)dy\right)dx + c_1\right]\exp\left(\int_{t_0}^t b_1(x)dx\right),$$

where $c_1 = u(t_0) = g(t)$, so that

$$\left|u(t)\cdot\exp\left(-\exp\left(-\int_{t_0}^t b_1(x)dx\right)\right)\right|$$
$$\leqq \int_{t_0}^t |b_2(x)|\,|u(x)|\,|u(x-\lambda)|\,\exp\left(-\int_{t_0}^x b_1(y)dy\right)dx + |c_1|.$$

Since $|u(t)| \leqq 1$, we have $|u(t)|\;|u(-\lambda)| \leqq |u(t-\lambda)| \leqq |u(t-\lambda)| + |u(t)|$; thus

$$|u(t-\lambda)| \leqq 1 + \frac{|u(t-\lambda)|}{|u(t)|}\cdot \tag{2}$$

Consequently,

$$\tag{3}\begin{aligned}\left|u(t)\exp\left(-\int_{t_0}^{t} b_1(x)dx\right)\right| &\leqq \int_{t_0}^{t} |b_2(x)|\,|u(x)|\exp\left(-\int_{t_0}^{x} b_1(y)dy\right)dx\\ &+\int_{t_0-\lambda}^{t_0} |b_2(x+\lambda)|\,|u(x)|\exp\left(-\int_{t_0}^{x+\lambda} b_1(y)dy\right)dx\\ &+\int_{t_0}^{t} |b_2(x+\lambda)|\,|u(x)|\exp\left[-\int_{t_0}^{x} b_1(y)dy-\int_{x}^{x+\lambda} b_1(y)dy\right]dx+|c_1|.\end{aligned}$$

Now we estimate the integral on the right side of (3). By assumption, there exists $0<M_1<\infty$, such that $\int_t^{t+\lambda} b_1(x)dx<M_1<\infty$. In the integral

$$\int_{t_0-\lambda}^{t_0} |b_2(x+\lambda)|\,|u(x)|\exp\left(-\int_{t_0}^{x+\lambda} b_1(y)dy\right)dx,$$

on replacing the function $u(t)$ by the initial function $g(t)$, we find $0<M_2<\infty$, such that

$$\max_{t_0-\lambda\leqq t\leqq t_0}\int_{t_0-\lambda}^{t_0} |b_2(x+\lambda)|\,|u(x)|\exp\left(-\int_{t_0}^{x+\lambda} b_1(y)dy\right)dx\leqq M_2<\infty.$$

From this we obtain the inequality

$$\tag{4}\begin{aligned}&\left|u(t)\exp\left(-\int_{t_0}^{t} b_1(x)dx\right)\right|\\ &\leqq M_3+\int_{t_0}^{t} |u(x)|\,[\,|b_2(x)|+M_1|b_2(x+\lambda)|\,]\exp\left(-\int_{t_0}^{x} b_1(y)dy\right)dx,\end{aligned}$$

where $M_3=M_2+|c_1|$. We employ Bellman's inequality in (4) to obtain

$$\left|u(t)\exp\left(-\int_{t_0}^{t} b_1(x)dx\right)\right|\leqq M_3\exp\left[\int_{t_0}^{t}(|b_2(x)|+M_1|b_2(x+\lambda)|)dx\right]$$

whence,

$$|u(t)|\leqq M_3\exp\left[\int_{t_0}^{t} b_1(x)dx+\int_{t_0}^{t}(|b_2(x)|+M_1|b_2(x+\lambda)|)dx\right].$$

Therefore, under the stated assumptions, the theorem is proved.

REMARK. (i) The theorem is still true even if $|u(t)|\leqq 1$ is not satisfied, provided that (2) is satisfied. (ii) If $|u(t)|\leqq C$ (C is constant) then the theorem is still valid. In fact, the inequality (2) is replaced by

$$|u(t-\lambda)|\leqq 1+\frac{C|u(t-\lambda)|}{|u(t)|}.$$

Finally, we prove the boundedness of all solutions of (1); namely,

THEOREM 2. *Let the following conditions be satisfied*

(i) *the trivial solution of the linear first approximation is asymptotically stable,*

(ii) $|u(t)| \leqq 1$ *or* $|u(t-\lambda)| \leqq 1+|u(t-\lambda)|/|u(t)|$ *for all* $t \geqq t_0$,

(iii) $\int_{t_0}^{t} |b_2(x+\lambda)| dx$ *and* $\int_{t_0}^{t} |F(x)| dx$ *are bounded, when* $t \geqq t_0$,

then all solutions of (1) *are bounded.*

The proof is similar to that of Theorem 1.

References

1. Chian Hod-Ming, The Asymptotic Stability of a Differential-Difference Equation with Lag, Journal of Shautung University, Math., 1 (1960) in Chinese.

2. A. L. Aliskolish, The stability of solutions of differential-difference equations, Progress of Math., 9.4, 64 (1954) in Russian.

3. ———, Differential Equations, (in Russian).

4. N. N. Krasovskii, Several problems concerning the stability of motion, 1959, in Russian.

5. A. D. Myshkis, Linear differential equations with lag, 1951, in Russian.

6. R. Bellman, Stability Theory of Differential Equations, McGraw-Hill, New York, 1953.

SOME EXPANSIVE HOMEOMORPHISMS OF THE REALS

B. F. BRYANT AND D. B. COLEMAN, Vanderbilt University

1. Introduction. A homeomorphism f of a metric space X onto X is *expansive* (with expansive constant $\delta > 0$) provided for each pair $x, y \in X$ with $x \neq y$, there exists an integer n such that $d(f^n(x), f^n(y)) > \delta$. If X is compact, then f is expansive if and only if $f^m (m \neq 0)$ is expansive [2]. Probably the first expansive homeomorphisms of the reals which come to mind are those defined by $f_k(x) = kx$, where $k \neq -1, 0, 1$; these homeomorphisms also have the property that f_k^m is expansive for each integer $m \neq 0$. Furthermore, since

$$d(f_k^n(x), f_k^n(y)) = |k|^n |x - y|,$$

each $\delta > 0$ is an expansive constant for f_k. If X is compact, however, then the expansive constants of an expansive homeomorphism are bounded and the least upper bound of the expansive constants is not an expansive constant [1]. Thus the following questions arise:

Is there an expansive homeomorphism f of the reals (or more generally, n-dimensional Euclidean space E^n) such that

(1) f^m *is not expansive if* $m \neq -1, 1$?

(2) *the expansive constants of* f *are bounded*?

The purpose of this paper is to give an affirmative answer to both of these questions.

We use homeomorphisms of the following type: let $A = [a_0, \infty)$ where $a_0 > 0$, and let $I_i = [a_i, a_{i+1}]$, where $a_i < a_{i+1}$ for $i = 0, 1, 2, \cdots$, be a countably infinite partitioning of A. Let f be a self-homeomorphism of $[0, \infty)$ such that

$$f[0, a_0] = [0, a_1]. \tag{1}$$

$$f(x) > x \qquad \text{for } 0 < x < a_0. \tag{2}$$

$$f(I_i) = I_{i+1} \quad \text{for } i = 0, 1, 2, 3, \cdots. \tag{3}$$

If $x>0$, then $f(x)>x$ and $f^n(x)\in I_0$ for some n. Also, if $0<x<y$, then there exists z such that $x<z<y$ and $f^n(x)$, $f^n(z)\in I_0$ for some n; since f is increasing, $|f^k(x)-f^k(z)|<|f^k(x)-f^k(y)|$ for each integer k. Hence, f is expansive (with expansive constant δ) if and only if $u, v\in I_0$, $u\neq v$, implies $|f^n(u)-f^n(v)|>\delta$ for some integer n. There is nothing special about I_0; any other I_i may be used.

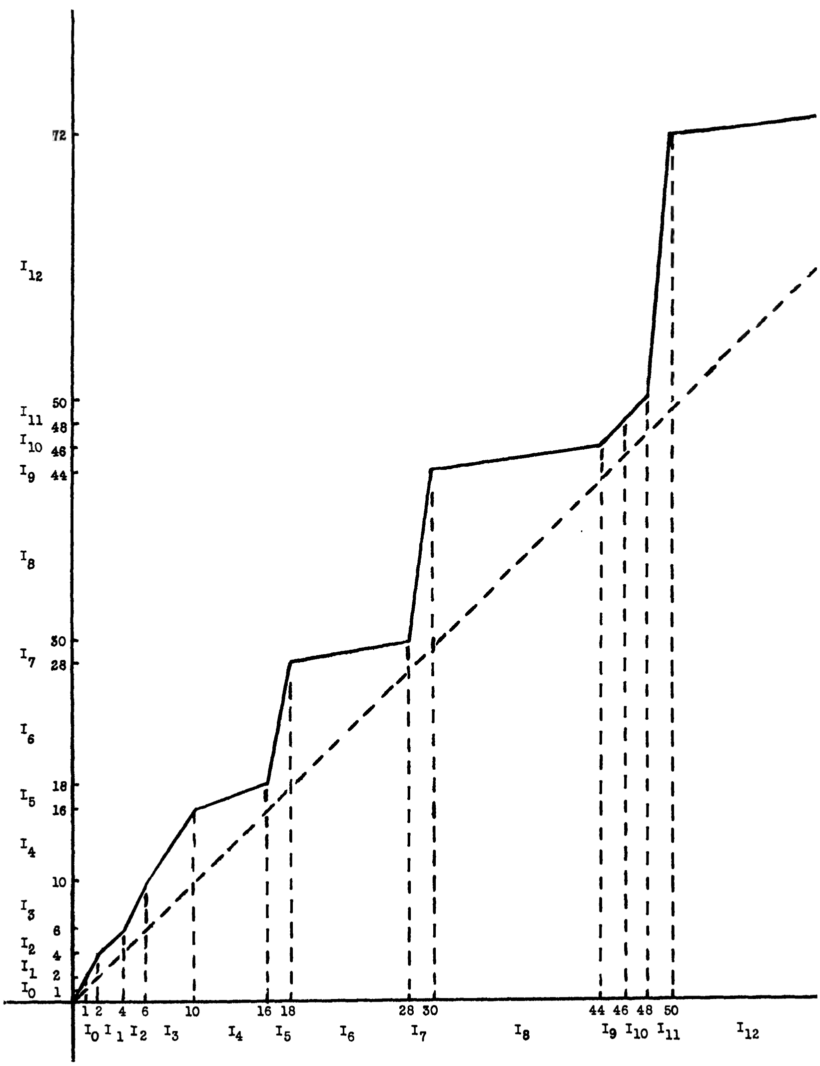


FIG. 1

2. The Examples.

Example 1. Let $I_0=[1, 2]$, $I_1=[2, 4]$, $f(x)=2x$ for $x\in[0, 2]$ (note that this defines f on both $[0, a_0]$ and $I_0=[a_0, a_1]$), and $f(x)=x+2$ for $x\in[2, 4]=I_1$. Define $m_1=1$ and $b_1=2$. For $x\in I_n$ $(n=2, 3, 4, \cdots)$, let $f(x)=m_nx+b_n$ where

$$m_nm_{n-1}\cdots m_2m_1=\begin{cases}1 & \text{if } n \text{ is not prime}\\ n & \text{if } n \text{ is prime}\end{cases}, \quad \text{and} \quad b_n=(m_{n-1}-m_n)a_n+b_{n-1}.$$

Then the graph of f is a sequence of straight line segments, and f is a homeomorphism of the type described in Section 1. A portion of the graph of f is given in Fig. 1.

If $x\in I_1$, then $f^n(x)=m_nm_{n-1}\cdots m_2m_1x+B_n$ where B_n is a constant. Let $\delta>0$, u, $v\in I_1$ with $u\neq v$. Select a prime p such that $p|u-v|>\delta$. Then

$$\begin{aligned}|f^p(u)-f^p(v)| &= |m_pm_{p-1}\cdots m_2m_1u+B_p-m_pm_{p-1}\cdots m_2m_1v-B_p|\\ &= |pu-pv| = p|u-v| > \delta.\end{aligned}$$

Thus f is expansive and each positive real number is an expansive constant. We now show that f^k is not expansive if $k>1$. Let $\delta>0$ and select u, $v\in I_1$ such that $|u-v|\leqq\delta/k$. Then

$$|f^k(u)-f^k(v)|=\begin{cases}k|u-v| & \text{if } k \text{ is prime}\\ |u-v| & \text{if } k \text{ is not prime}\end{cases}\leqq\delta.$$

If $i>1$, then ki is not prime, and hence $|f^{ki}(u)-f^{ki}(v)|=|u-v|<\delta$. If $i<0$, then $f^{ki}(u)$ and $f^{ki}(v)$ are in $[0, 2]$ and $|f^{ki}(u)-f^{ki}(v)|\leqq|u-v|<\delta$. Thus f^k is not expansive. Since f^k is expansive if and only if f^{-k} is expansive, we have established that f^k is not expansive if $k\neq-1, 1$.

Using this homeomorphism on each ray emanating from the origin yields a self-homeomorphism h of E^n such that h^k is expansive if and only if $k=-1, 1$, and for which each $\delta>0$ is an expansive constant.

Example 2. Let $I_n=[n+1, n+2]$, $n=0, 1, 2, \cdots$. Select a countable dense subset A of $I_0=[1, 2]$, and let $B=\{(c, d): (c, d)\in A\times A \text{ and } c\neq d\}$. Let (c_i, d_i), $i=0, 1, 2, \cdots$, be the members of B, and let $0<\delta<1$. For each nonnegative integer n, let ϕ_n be a homeomorphism of I_n onto I_{n+1} such that $\phi(n+1)=n+2$, $\phi(n+2)=n+3$ and $|\phi_n\phi_{n-1}\cdots\phi_0(c_n)-\phi_n\phi_{n-1}\cdots\phi_0(d_n)|>\delta$. Define

$$f(x)=\begin{cases}2x & \text{if } 0\leqq x\leqq 1.\\ \phi_n(x) & \text{if } x\in I_n.\end{cases}$$

Then f is a homeomorphism of the type described in Section 1. If $x\in I_0$, then $f^n(x)=\phi_n\phi_{n-1}\cdots\phi_0(x)$. For x, $y\in I_0$, with $x<y$, select $(c_n, d_n)\in B$ such that $x<c_n<d_n<y$. Then

$$\begin{aligned}|f^n(x)-f^n(y)| &> |f^n(c_n)-f^n(d_n)|\\ &= |\phi_n\phi_{n-1}\cdots\phi_0(c_n)-\phi_n\phi_{n-1}\cdots\phi_0(d_n)| > \delta.\end{aligned}$$

Thus f is expansive and δ is an expansive constant. But $u, v \in I_0$ implies $|f^n(u)-f^n(v)| \leqq 1$ for each n; thus 1 is not an expansive constant.

Again, this homeomorphism may be used on each ray emanating from the origin to obtain an expansive self-homeomorphism of E^n with bounded expansive constants.

References

1. B. F. Bryant, Expansive self-homeomorphisms of a compact metric space, this MONTHLY, 69 (1962) 386–391.
2. W. R. Utz, Unstable homeomorphisms, Proc. Amer. Math. Soc., 1 (1950) 769–774.

BALANCED FIELD EXTENSIONS

JOSEPH LIPMAN, Purdue University

Let k be a field, and let K be an algebraic extension of k. K is then a purely inseparable extension of a separable extension of k; for reasons of symmetry, one might wonder when K will be a separable extension of a purely inseparable extension of k. (This is not always so: cf. example at the end of this note.) When this does happen, let us say that K/k is "balanced." We wish to set down some simple observations about such extensions.

For basic notions of field theory see O. Zariski and P. Samuel: *Commutative Algebra*, Van Nostrand, Princeton, N. J., 1958, Volume I, chapter 2; also chapter 3, section 15, for the definition and properties of free joins.

PROPOSITIONS: A. *The following are equivalent*:

1. K/k *is balanced.*
2. *There exists a separable algebraic extension of* K *which is normal over* k.
3. *If* L *is a field of algebraic functions over* k, *then the order of inseparability* $[(L, K): K]_i$ *is the same for all free joins* (L, K) *of* L/k *and* K/k.

In geometric language, 3. reads: If V/k is an irreducible algebraic variety, then all the irreducible components of V/K have the same order of inseparability.

B. *Let* $\bar{k}$ *be an algebraic closure of* k, *and let* $\bar{k}_s$ (*respectively* $\bar{k}_i$) *be the subfield of* $\bar{k}$ *consisting of all elements which are separable,* (*resp. purely inseparable*) *over* k. *We know that the subfields of* $\bar{k}$ (*resp.* $\bar{k}_s$, $\bar{k}_i$) *form a lattice* Z (*resp.* Z_s, Z_i) *under the operations of field composition and intersection.*

The balanced extensions of k *in* $\bar{k}$ *form a sublattice of* Z *isomorphic with the direct product* $Z_s \times Z_i$.

Proofs: A. We show that $1 \xrightarrow{a)} 3 \xrightarrow{b)} 2 \xrightarrow{c)} 1$.

a) Let $k \subseteq I \subseteq K$, I being a field such that I/k is purely inseparable, and K/I is separable. Any free join (L, K) contains a free join (L, I), and since K/I is separable, we have $[(L, K): K]_i = [(L, I): I]_i$. Thus we may assume that

$K=I$. But then there is nothing to prove, since all free joins of L/k and I/k are equivalent.

b) Let $x \in K$, and let $L=k(x)$. The free joins of L/k, K/k, are all the fields of the form $K(\bar{x})$, where $\bar{x}$ is k-conjugate to x, in some fixed algebraic closure $\overline{K}$ of K. Since $[K(x): K]_i=[K: K]_i=1$, we have $[K(\bar{x}): K]_i=1$, i.e. *all k-conjugates of x (in $\overline{K}$) are separable over K.* From this it follows immediately that if N is the least extension of K normal over k, then N/K is separable.

c) Let $N \supseteq K$ be such that N/K is separable and N/k is normal. Let $I \subseteq N$ be the field of invariants of all automorphisms of N/k. Then I/k is purely inseparable, and N/I is separable. It will be sufficient to show that $I \subseteq K$. But any x in I is separable over K (since N/K is separable), and purely inseparable over K (since I/k is purely inseparable).

B. Let B be the set of balanced extensions of k in $\bar{k}$. Clearly $K \in B$ iff $K/(K \cap \bar{k}_i)$ is separable, and this latter condition may be expressed as follows:

(1′) If $x \in K$, if $f(X)$ is the minimum monic polynomial of x over k, and if $\bar{f}(X) \in \bar{k}[X]$ is *the* polynomial without multiple roots, of which $f(X)$ is a power, then $\bar{f}(X) \in (K \cap \bar{k}_i)[X]$.

It follows immediately that an arbitrary intersection of members of B is again a member of B.

Again, if $K \in B$, then K is a separable extension of $K \cap \bar{k}_i$; also K is purely inseparable over $K \cap \bar{k}_s$; if K' is the compositum $(K \cap \bar{k}_s, K \cap \bar{k}_i)$, then K/K' is both separable and purely inseparable, i.e., $K=K'$. We see then, that *$K \in B$ iff K is generated by a separable extension of k and a purely inseparable extension of k.*

It follows easily that the field generated by an arbitrary collection of members of B is itself a member of B.

We have shown, therefore, that B is a sublattice of Z (in fact, a complete sublattice). We have also given an order-preserving map F from $Z_s \times Z_i$ *onto* B; if $S \in Z_s$, $I \in Z_i$, then $F(S, I)$ is the composed field (S, I). Now if $x \in (S, I) \cap \bar{k}_i$, then x is separable over I and purely inseparable over I; hence $(S, I) \cap \bar{k}_i=I$. Similarly $(S, I) \cap \bar{k}_s=S$. Thus F is injective, and the proof is complete.

EXAMPLE. For an example of a nonbalanced extension, let L be a field of characteristic two, let Y, Z, be indeterminates over L, let $k=L(Y, Z)$, and let $K=k(x)$, x being a root of

$$f(X) = X^4 + YX^2 + Z = 0.$$

One checks that $f(X)$ is irreducible over k, that $[K: k]_i=2$, and that, in the notation of (1′) above, $\bar{f}(X)=X^2+\sqrt{Y}X+\sqrt{Z}$. According to (1′), K/k cannot be balanced unless $\sqrt{Y} \in K$ and $\sqrt{Z} \in K$. Since $[k(\sqrt{Y}, \sqrt{Z}): k]_i=4$, this is impossible.

REPRESENTATIONS OF INTEGERS PRIME TO A GIVEN INTEGER

J. P. WILSON, University College of Wales, Aberystwyth, Wales

The following theorem is implicit in the calculation of the number of modular substitutions incongruent (mod n), (see e.g. [1], p. 391) and is well known.

THEOREM 1. *Let n be a positive integer and a, b integers such that $(a, b, n)=1$. Then there exist integers $a' \equiv a$, $b' \equiv b$* (mod n) *such that $(a', b')=1$.*

The object of this note is to prove the following extension of this theorem.

THEOREM 2. *Let n be a positive integer and a, b integers such that $(a, b, n)=1$, and let d be any positive integer such that $(d, n)=1$. Then there exist integers $a'' \equiv a$, $b'' \equiv b$* (mod n) *such that*

$$(a'', b'') = d.$$

For completeness, since we use Theorem 1 in the proof of Theorem 2, we give a proof of Theorem 1.

Proof of Theorem 1. Choose any nonzero $a' \equiv a$ (mod n). Then

$$(a', b, n) = (a, b, n) = 1.$$

Let p_i, $i=1, \cdots, r$ be the primes which are divisors of a', but not of b and not of n, and let $\alpha = \Pi_{i=1}^{r} p_i$. (If $r=0$, then $\alpha=1$.) Then $(a', b+\alpha n)=1$, i.e. we may take $b'=b+\alpha n$.

NOTE. This result also follows from Dirichlet's theorem that if $(a, d)=1$, then there are infinitely many primes of the form $a+kd$.

Proof of Theorem 2. Since $(d, n)=1$, $(d, n) | a, b$ and there exist integers k, l, ρ, σ such that

$$\begin{aligned} -kn + \rho d &= a \quad \text{and} \\ -ln + \sigma d &= b. \end{aligned} \tag{1}$$

Then

$$(a + kn, b + ln) = (\rho d, \sigma d) = (\rho, \sigma)d,$$

and hence, if $(\rho, \sigma)=1$, we may take $a''=a+kn$, $b''=b+ln$. Otherwise we note that $(\rho, \sigma, n) | a, b$ by (1) and hence $(\rho, \sigma, n) | (a, b, n)=1$. Then by Theorem 1 there exist integers $\rho' \equiv \rho$, $\sigma' \equiv \sigma$ (mod n) such that $(\rho', \sigma')=1$. Let $a''=\rho' d \equiv \rho d \equiv a$ (mod n) and $b''=\sigma' d \equiv b$ (mod n). Then

$$(a'', b'') = (\rho' d, \sigma' d) = (\rho', \sigma')d = d.$$

COROLLARY. *Let A and B be residue classes* (mod n) *such that, if $a \in A$ and $b \in B$, $(a, b, n)=1$. Then the sets*

$$S = \{(a, b); a \in A, b \in B\}, \qquad S' = \{d; d > 0, (d, n) = 1\}$$

are identical.

Proof. By the theorem $S' \subset S$. But trivially $S \subset S'$, since, if $a \in A$, $b \in B$ and $(a, b) = d$, then $(d, n) = ((a, b), n) = (a, b, n) = 1$.

I am indebted to the referee for two helpful suggestions.

Reference

1. F. Klein and R. Fricke, Theorie der elliptischen Modulfunctionen, I, Teubner, 1890.

ORDERS FOR FINITE NONCOMMUTATIVE RINGS

D. B. Erickson, Concordia College, Moorhead, Minnesota

Introductory Comments. The solution to problem E 1529 (this Monthly, 70, 1963, p. 441) answers the question of the order for the smallest noncommutative ring. It is 4. This paper answers the more general question concerning a characterization of the orders for finite noncommutative rings in a series of three theorems. (In view of the first theorem the comment following the solution of E 1529, to the effect that there exists a noncommutative ring of each composite order N, must be incorrect.)

Theorem 1. *If R is a finite ring of order $n > 1$ and if n has square free factorization, then R is a commutative ring.*

Proof. It is well known that if the additive group of a ring is cyclic then the ring is commutative. The additive group of R is a finite abelian group which, by the Basis Theorem for Finite Abelian Groups, may be expressed as the direct product of s cyclic subgroups of R, say $R_1, \cdots, R_s$, where the order of each R_i divides the order of R_{i+1} $(i = 1, \cdots, s-1)$, and s is the number of elements in a minimal generating system (basis) for R. But by the Lagrange Theorem the order of each R_i divides the order of R. Since the order of R is square free, the additive group of R is the cyclic group of order n (abbr. C_n), and therefore R is a commutative ring.

In the next two theorems we construct a noncommutative ring of order m for each positive integer $m > 1$ having square factors.

Theorem 2. *If p is a prime integer, then there exists a noncommutative ring of order p^2.*

Proof. Let R be the direct product of C_p with itself. It is clear that R is an abelian group of order p^2 and that R is not cyclic. A minimal generating system for R is $\{(a, 0), (0, a)\}$, where a is a generator for C_p. We define multiplication on the basis elements by the requirement that the product of two basis elements is the left factor, and extend it to the whole system by the distributive law, whence

$$(j_1a, k_1a) \cdot (j_2a, k_2a) = (j_2 + k_2)(j_1a, k_1a),$$

for all (j_ia, k_ia) in R, where the j_i and k_i are integers. Closure, the associative

law, and the distributive laws are easily verified; thus, R is a ring of order p^2. Since $(0, a)(a, 0) = (0, a)$ but $(a, 0)(0, a) = (a, 0)$ R is not commutative.

This last result is extended to all orders np^2 by Theorem 3.

THEOREM 3. *Let R_1 be a ring of order p^2 as constructed in Theorem 2, and let R_2 be any ring of order n, then the ring $R = R_1 \dotplus R_2$ (i.e. the direct sum of R_1 and R_2) is a noncommutative ring of order np^2.*

REMARK. The proof of this statement is obvious from the properties of direct sums of rings. We only remark that for R_2 we may use the trivial ring of order n which has as its additive group C_n and in which all products are zero; note also that R contains a subring isomorphic to R_1 and is therefore noncommutative.

We summarize in the following corollary.

COROLLARY. *If m is a positive integer, $m > 1$, then there exists a noncommutative ring of order m if and only if m has square factors.*

Supported by NSF Grant GE-1082.

ON SOME THEOREMS OF BRAM FOR SUBNORMAL OPERATORS

G. H. ORLAND, University of Illinois, Urbana

Subnormal operators were defined and studied by Halmos in [3] and [4]. His results were further extended by Bram [1]. A (bounded) operator S on a complex Hilbert space H is *subnormal* if on a Hilbert space K containing H as a subspace, there is a normal operator N whose restriction to H is S. In other words, an operator is subnormal if it is the restriction of a normal operator to an invariant subspace. To avoid trivial complications it is customary to assume that N is a minimal extension of S in the sense that the only subspace of K which contains H and on which N is normal is K itself. For the spectra of S and N, the interesting result, $\sigma(S) \supset \sigma(N)$, was proved in [4]. The resolvent set of N can be expressed as

$$\rho(N) = U_\infty \cup \bigcup_{n=1}^{\infty} U_n,$$

where each U is one of its (open) connected components; U_∞ is the single unbounded component. Following Bram, each bounded U_n is called a *hole* of $\sigma(N)$. Set $H(N) = \bigcup_{n=1}^{\infty} U_n$. In Theorem 3 of [1], Bram proved that $\sigma(S) \subset \sigma(N) \cup H(N)$. In Theorem 4 he proved that for any hole of $\sigma(N)$, either all of it is contained in $\sigma(S)$, or none of it. By using some ideas from [5], a simpler and simultaneous proof of both theorems will be given.

Let $d(\lambda, \sigma(A))$ represent the distance, in the complex plane, from the point λ to $\sigma(A)$. Then from $\sigma(S) \supset \sigma(N)$ it follows that $d(\lambda, \sigma(S)) \leq d(\lambda, \sigma(N))$ for any λ. If $\lambda \in \rho(S)$ so that the resolvents $R_\lambda(S)$ and $R_\lambda(N)$ both exist, then $R_\lambda(N)$ extends $R_\lambda(S)$ and $\|R_\lambda(N)\| \geq \|R_\lambda(S)\|$. It is also true that for any operator A,

$[d(\lambda, \sigma(A))]^{-1} \leqq \|R_\lambda(A)\|$ (cf. [2, p. 566]). When in addition A is normal, the spectral theorem yields

$$\|R_\lambda(A)x\|^2 = \int \frac{d(E_\mu x, x)}{|\mu - \lambda|^2} \leqq \int \frac{d(E_\mu x, x)}{[d(\lambda, \sigma(A))]^2} = \left[\frac{\|x\|}{d(\lambda, \sigma(A))}\right]^2,$$

so $\|R_\lambda(A)\| = [d(\lambda, \sigma(A))]^{-1}$. Putting all this together gives

$$[d(\lambda, \sigma(S))]^{-1} \leqq \|R_\lambda(S)\| \leqq \|R_\lambda(N)\| = [d(\lambda, \sigma(N))]^{-1} \leqq [d(\lambda, \sigma(S))]^{-1}.$$

Therefore $d(\lambda, \sigma(S)) = d(\lambda, \sigma(N))$.

Now suppose there are points of $\sigma(S)$ in some U but that $U \not\subset \sigma(S)$. (In particular $U_\infty \not\subset \sigma(S)$.) Then there is a boundary point μ of $\sigma(S)$ in that U. Therefore a point in $U \cap \rho(S)$ can be found which is closer to μ (and consequently to $\sigma(S)$) than it is to $\sigma(N)$. Since this is impossible, the theorems are proved.

References

1. J. Bram, Subnormal operators, Duke Math. J., 22 (1955) 75–94.
2. N. Dunford and J. T. Schwartz, Linear operators I, Interscience, New York, 1958.
3. P. R. Halmos, Normal dilations and extensions of operators, Summa Brasil. Math., 2 (1950) 125–134.
4. ———, Spectra and spectral manifolds, Ann. Soc. Polon. Math., 25 (1952) 43–49.
5. G. H. Orland, On a class of operators, Proc. Amer. Math. Soc., 15 (1964) 75–79.

A PROPERTY OF FUNCTIONS DISCONTINUOUS ON A DENSE SET

G. A. HEUER, Concordia College, Moorhead, Minnesota

Let f be a function on the reals to the reals. Let D, L, $H(\alpha)$, and C be, respectively, the sets of points where f is differentiable, f is Lipschitzian, f satisfies a Hölder condition of order $\alpha(0<\alpha\leqq 1)$, and f is continuous. If $0<\alpha_1<\alpha_2<1$, then $D \subseteq L = H(1) \subseteq H(\alpha_2) \subseteq H(\alpha_1) \subseteq C$. This note is concerned with a simultaneous extension of the following two theorems.

THEOREM 1. (Fort, [2]) *If C' (the complement of C) is dense, then D is a set of the first category.*

THEOREM 2 (Heuer et al., [3]). *If C is the set of irrational numbers, then L' is uncountable and dense.*

Recall also the following theorem of classical analysis [1, p. 102].

THEOREM 3. *If C is dense, then C' is of first category.*

It is a corollary of Theorems 1 and 3 that if both C and C' are dense, then $C \cap D'$ is a residual set. It follows that D' is uncountable and dense. This bears some similarity to Theorem 2, but neither implies it nor is implied by it.

However, by making only slight modifications in the proof of Fort's theo-

COMPLETION OF PRIMITIVE MATRICES

IRVING REINER, University of Illinois, Urbana

Let R be any ring in which classical ideal theory is valid (see [1]). For example, R might be chosen as the ring of all algebraic integers in some algebraic number field. A square matrix with entries in R, whose determinant is a unit in R, is called *unimodular*. An $r\times n$ matrix A, where $r\leqq n$, is said to be *completable* if there exists an $n\times n$ unimodular matrix of which A forms the first r rows. Which matrices are completable?

If A is the $r\times n$ matrix consisting of the first r rows of the unimodular $n\times n$ matrix U, then the determinant of U is expressible as a linear combination of the $r\times r$ minors of A. Therefore the $r\times r$ minors of A generate the unit ideal of R, and we describe this situation by saying that A is *primitive*. Thus, completable matrices are necessarily primitive. We note that if A is primitive, and T is unimodular, then AT is also primitive.

By using the theory of elementary divisors, Steinitz [2] showed, conversely, that every primitive matrix is completable. In 1956, I gave a simple proof [3] for the following special case:

A row vector is completable if and only if it is primitive.

In a recent note [4], S. Ram gave an incorrect proof of the general result. It therefore seems worthwhile to indicate how the general result follows readily from the above special case.

Let A be a primitive $r\times n$ matrix, where $n\geqq r\geqq 1$, and assume that every primitive $(r-1)\times(n-1)$ matrix is completable. Since A is primitive, so is its first row $\mathfrak{a}$. By the special case, $\mathfrak{a}$ is completable, and thus there exists a unimodular $n\times n$ matrix X such that

$$\mathfrak{a}X = (1\ 0\ \cdots\ 0).$$

Therefore we may write

$$AX = \begin{bmatrix} 1 & 0\ \cdots\ 0 \\ * & \\ \cdot & \\ \cdot & A_1 \\ \cdot & \\ * & \end{bmatrix}.$$

Now AX is also primitive, since A is. Further, the nonzero $r\times r$ minors of AX coincide with the nonzero $(r-1)\times(r-1)$ minors of A_1. Therefore A_1 is a primitive $(r-1)\times(n-1)$ matrix, so by the induction hypothesis there exists a matrix B_1 for which

$$\begin{bmatrix} A_1 \\ B_1 \end{bmatrix}$$

is unimodular. Let

$$Y = \begin{bmatrix} 1 & 0 \cdots 0 \\ * & \\ \cdot & A_1 \\ \cdot & \\ * & \\ 0 & \\ \cdot & B_1 \\ \cdot & \\ 0 & \end{bmatrix}$$

Then Y is also unimodular, and A consists of the first r rows of YX^{-1}. This shows that A is completable, and finishes the proof of the general result.

References

1. B. L. van der Waerden, Modern Algebra II, Ungar, New York, 1950, p. 83.
2. E. Steinitz, Rechteckige Systeme und Moduln in algebraischen Zahlenkörpern, Math. Ann., 71 (1911) 328–354.
3. I. Reiner, Unimodular complements, this MONTHLY, 63 (1956) 246–247.
4. S. Ram, On the completion of rectangular matrices to unimodular matrices, Ganita, 13 (1962) 37–39.

IMBEDDING A SKELETON OF A SIMPLEX IN EUCLIDEAN SPACE

J. L. CHRISLOCK, University of California at Davis

It is well known that an n-dimensional complex K can be imbedded in R^{2n+1} ($(2n+1)$-dimensional Euclidean space). Flores [1] proved that the n-dimensional skeleton of the $(2n+2)$-dimensional simplex s^{2n+2} cannot be imbedded in R^{2n}. In this note we will determine the Euclidean space of smallest dimension in which the k-dimensional skeleton of the n-dimensional simplex can be imbedded.

If K is imbeddable in R^m but not R^{m-1} let $\phi(K)=m$; $\phi(r, s)$ will denote $\phi((s^r)^s)$, where $(s^r)^s$ is the s-dimensional skeleton of the r-dimensional simplex.

LEMMA 1. *If $\phi(r, s) \leqq n$ then $\phi(r+1, s) \leqq n+1$.*

Proof. Let K be a complex isomorphic to $(s^r)^s$ in $R^n \subseteq R^{n+1}$. Pick a point v in R^{n+1} which is not in R^n. Form the join of v with each $(s-1)$-dimensional face of K. The join considered as an abstract complex will be isomorphic to the complex $(s^{r+1})^s$.

LEMMA 2. $\phi(n+2, n) \leqq n+1$.

Proof. Since $(s^{n+1})^n$ is isomorphic to the n-sphere, $\phi(n+1, n)=n+1$. Let b_s be the barycenter of s^{n+1}, and form the join of b_s with each of the $(n-1)$-dimensional faces of $(s^{n+1})^n$. The resulting complex will contain all possible subsets

with $n+1$ elements of the set consisting of all vertices of s^{n+1} with b_s adjoined. Hence the join will be isomorphic to $(s^{n+2})^n$.

THEOREM. *For any positive integer* n, $\phi(n, n)=n$, $\phi(n+1, n)=n+1$, $\phi(n+k, n)=n+k-1$ *for* $k=2, 3, \cdots, n+1$, *and* $\phi(n+k, n)=2n+1$ *for* $k\geqq n+2$.

Proof. Clearly $\phi(n, n)=n$, and since $(s^{n+1})^n$ is homeomorphic to the n-sphere, $\phi(n+1, n)=n+1$. Now $(s^{n+2})^n$ has as a subcomplex $(s^{n+1})^n$. By Lemma 2 this tells us that $\phi(n+2, n)=n+1$. Using Lemma 1 repeatedly we get from $\phi(n+2, n)=n+1$ that

$$\phi(n+2+j, n) \leqq n+1+j \quad \text{for } j=1, 2, \cdots, n-1.$$

Assume that, for some $j_0\leqq n-1$, $\phi(n+2+j_0, n)\leqq n+j_0$. This would imply that $\phi(n+2+n, n)=\phi(2n+2, n)\leqq 2n$, contradicting Flores' result. Hence $\phi(n+2+j, n)=n+1+j$ for $j=1, 2, \cdots, n-1$. Finally Flores' result tells us that $\phi(n+k, n)=2n+1$ for $k\geqq n+2$. Thus the theorem is proved.

Reference

1. G. Flores, Über n-dimensionale Komplexe die im R_{2n+1} absolut selbstverschlungen sind, Ergebnisse eines mathematischen Kolloquiums, 6 (1934) 4–7.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

DETERMINANTAL REPRESENTATIONS OF RECURSION RELATIONS FOR CERTAIN ORTHOGONAL POLYNOMIALS

F. M. STEIN, Colorado State University, M. S. HENRY, Colorado State University, and L. G. KING, Pacific Northwest Laboratories, Battelle Memorial Institute

Standard derivations of the three term recursion relations connecting orthogonal polynomials $P_n(x)$, $P_{n-1}(x)$, and $P_{n+1}(x)$ are well known, see [**1**, **2**, **6**]. In this note we observe that these recursion relations for four types of orthogonal polynomials may be derived from the determinant representation of orthogonal polynomials of Pandres [**4**, **5**] and may be expressed in terms of these determinants. Also shown is the connection between the various polynomials and their determinantal representation so that the usual form of the recursion relations can be readily obtained.

1. Properties of orthogonal polynomials. It is known (see [**1**]) that for each of the sets of classical orthogonal polynomials, there exists a generalized Rodrigues formula by which the nth member, P_n, is given. This formula may be expressed as

$$K_nP_n = \frac{1}{w(x)}\frac{d^n}{dx^n}\left[w(x)F^n\right], \tag{1}$$

where K_n is a constant; a particular set of polynomials is determined by the choice of the functions $w(x)$ and $F(x)$.

It was shown by Pandres [4, 5] that the nth derivative of a function $f(x)$ may be given by

$$\frac{d^n}{dx^n}\left[f(x)\right] = f(x)\Delta_n, \tag{2}$$

where the symbol Δ_n denotes the nth order determinant

$$\Delta_n = \begin{vmatrix} D_1 & -1 & 0 & \cdots & 0 & 0 \\ D_2 & D_1 & -2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ D_{n-1} & D_{n-2} & D_{n-3} & \cdots & D_1 & 1-n \\ D_n & D_{n-1} & D_{n-2} & \cdots & D_2 & D_1 \end{vmatrix}. \tag{3}$$

Applying (2) to (1) and simplifying, we obtain

$$K_nP_n = \Delta_n, \tag{4}$$

in which $D_1, D_2, \cdots, D_n$ are defined by

$$D_k = \frac{F^k}{(k-1)!}\frac{d^k}{dx^k}\left[\log w + n\log F\right].$$

Using these results, Pandres [4, 5] showed that the Hermite, Laguerre, and Jacobi polynomials can be represented by Δ_n.

2. The classical polynomials of Hermite, Laguerre, and Jacobi. The Hermite polynomials are represented by Δ_n when $D_1 = -2x$, $D_2 = -2$, and $D_k = 0$ for $k > 2$. Using these values for D_k in (3) and expanding Δ_n by cofactors of the last column, we obtain

$$\Delta_n = -2x\Delta_{n-1} - (1-n)\begin{vmatrix} -2x & -1 & 0 & \cdots & 0 & 0 \\ -2 & -2x & -2 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & -2x & 2-n \\ 0 & 0 & 0 & \cdots & 0 & -2 \end{vmatrix}.$$

Upon expanding the last determinant by cofactors of the last row and then replacing n by $n+1$ in the result, we have

$$\Delta_{n+1} + 2x\Delta_n + 2n\Delta_{n-1} = 0, \tag{5}$$

a determinantal recursion relation for Hermite polynomials. Upon comparing the Rodrigues formula for Hermite polynomials with (1), we note that $K_n=1$. Hence $P_n=\Delta_n$ in (4), and we can thus obtain the standard recursion relation for Hermite polynomials from (5), namely,

$$H_{n+1}(x)+2xH_n(x)+2nH_{n-1}(x)=0.$$

For Laguerre polynomials $D_1=-x+c+n$, and $D_k=(-1)^{k-1}(n+c)$ for $k>1$ in (3). By proceeding in a manner similar to that for Hermite polynomials, we obtain the determinantal recursion relation

$$\Delta_n=(2n-1+c-x)\Delta_{n-1}-(n-1)(n-1+c)\Delta_{n-2}. \tag{6}$$

In this case $K_n=n!$ in (1), and thus $n!P_n(x)=\Delta_n$ in (4). Upon substituting into (6) we obtain the standard recursion relation for Laguerre polynomials,

$$nL_n(x)-(2n-1+c-x)L_{n-1}(x)+(n-1+c)L_{n-2}(x)=0.$$

For the Jacobi polynomials we have

$$D_k=-\left[(n+a)(x+1)^k+(n+b)(x-1)^k\right] \qquad \text{in (3).}$$

The determinantal recursion relation in this case is

$$\begin{aligned}(a+b+2n)&(a+b+2n+1)(a+b+2n+2)x\Delta_n\\ &=-(a+b+n+1)(a+b+2n)\Delta_{n+1}+(b^2-a^2)(a+b+2n+1)\Delta_n\\ &\quad-4n(a+n)(b+n)(a+b+2n+2)\Delta_{n-1}.\end{aligned} \tag{7}$$

The K_n of (1) is $2^n n!/(-1)^n$ and thus

$$\frac{2^n n!}{(-1)^n}P_n(x)=\Delta_n.$$

Upon substituting this result into (7) we get the standard recursion relation for Jacobi polynomials.

3. The generalized Bessel polynomials. A recently obtained type of orthogonal polynomials are the generalized Bessel polynomials [**3**]. These polynomials are orthogonal in a sense that is different from that for the preceding types of polynomials; here the integration is around a unit circle in the complex plane.

Upon applying the above procedures to the generalized Bessel polynomials we find that

$$D_k=(-1)^{k-1}x^{k-1}[(2n+a-2)x+kb].$$

The determinantal recursion relation is

$$(n+a-1)(2n+a-2)\Delta_{n+1}$$

$$(8) \qquad = \left[(2n+a)(2n+a-2)\frac{x}{b}+a-2\right](2n+a-1)b\Delta_n + n(2n+a)b^2\Delta_{n-1},$$

which may be converted to the three term recursion relation for Bessel polynomials [3] by noting that $K_n = b^n$.

The work of the authors F. M. Stein and M. S. Henry was supported by AFOSR Grant No. 185-63.

References

1. Bateman Manuscript Project. Higher Transcendental Functions, Vol. II. McGraw-Hill, New York, 1953.
2. D. Jackson, Fourier Series and Orthogonal Polynomials, MAA Carus Mathematical Monograph No. 6, 1941.
3. H. L. Krall, and Orrin Frink, A new class of orthogonal polynomials; the Bessel polynomials, Trans. Amer. Math. Soc., 65 (1949) 100–115.
4. D. Pandres Jr., On higher order differentiation, this MONTHLY, 64 (1957) 566–572.
5. ———, A determinant representation for the classical orthogonal polynomials, this MONTHLY, 67 (1960) 658–659.
6. G. Szegö, Orthogonal Polynomials, AMS Colloq. Publ., 23 (1939).

GENERALIZED BIRTHDAY PROBLEM

E. H. MCKINNEY, Ball State University

Introduction. The well-known "birthday problem" appears in most textbooks on probability; for instance see [1, pgs. 31–32]. The purpose of this paper is to generalize the problem and arrive at a solution which is applicable to a wider range of problems which are abstractly equivalent to the scheme of placing n balls in M cells. Many practical applications of such a scheme can be generalized from the situations described in [1, pgs. 10–11].

PROBLEM 1. *n people are selected at random. What is the probability that at least r people will have the same birthday?*

PROBLEM 2. *What is the smallest value of n such that the probability is greater than or equal to $1/2$ that at least r people have the same birthday?*

The solutions will be presented in a manner applicable to the wider range of similar problems referred to above.

Solution. Let X_i $(i=1, 2, 3, \cdots, n)$ be independent, identically distributed random variables with common distributions.

$$(1) \qquad \operatorname{Prob}[X_i = M_j \mid M_j \ (1 \leqq j \leqq M) \text{ is an integer}] = \frac{1}{M} \cdot$$

The probability that r or more X_i's are equal is to be determined. Let the event E be defined as "no r of the random variables X_i are equal." Then

Prob(r or more X_i's are equal) $=1-\text{Prob}(E)$.

Prob(E) is then computed by summing the probabilities of all ways in which n random variables can take on less than r equal values.

For a given n, let

n_1 = number of nonrepeated X_i's

n_2 = number of pairs of equal X_i's

n_3 = number of triples of equal X_i's

. . . .

n_{r-1} = number of $(r-1)$-tuples of equal X_i's, where obviously

$$n = \sum_{i=1}^{r-1} i n_i. \tag{2}$$

The general term of the summation representing Prob(E) is the probability that there are exactly n_1 nonrepeated items, n_2 pairs, n_3 triples, $\cdots$, $n_{(r-1)}$ $(r-1)$-tuples of equal X_i's which takes the form

$$\text{Prob}(n; n_1, n_2, \cdots, n_{r-1}) = \frac{n!}{\prod_{j=1}^{r-1} (n_j!)(j!)^{n_j}} \cdot \frac{\text{Perm}\left(M, \sum_{i=1}^{r-1} n_i\right)}{M^n}, \tag{3}$$

where Perm(a, b) is the number of permutations of a things taken b at a time.

The second factor of (3) is the probability that n independent, identically distributed random variables with common distribution (1) will have n_1 nonrepeated items, n_2 pairs, n_3 triples, $\cdots$, $n_{(r-1)}$ $(r-1)$-tuples of equal X_i's in a specified order. The first factor in (3) represents the number of distinguishable ways this particular order can be permuted.

It is interesting to note that one might mistakenly assume that the first factor in (3) should be the multinomial coefficient,

$$\left\{\begin{matrix} & & n & & \\ \underbrace{1, 1, \cdots, 1}_{n_1}, & \underbrace{2, 2, \cdots, 2}_{n_2}, & \cdots, & \underbrace{r-1, r-1, \cdots, r-1}_{n_{(r-1)}} \end{matrix}\right\}. \tag{4}$$

But this assumption counts as a different arrangement those permutations which interchange total groups of like multiple equal random variables. For example, if one has n random variables of which there are n_k k-tuples of equal random variables, interchanging the value taken on by the random variables in one k-tuple group with that of another k-tuple group does not produce a different permutation of n random variables with exactly n_k k-tuples of equal

random variables. Hence, one must divide the usual multinomial coefficient (4) by $\prod_{j=1}^{r-1} n_j!$, the number of ways these kinds of permutations can occur.

Finally then,

$$\text{(5)}\quad \text{Prob}(r \text{ or more } X_i\text{'s are equal}) = 1 - \sum \text{Prob}(n; n_1, n_2, n_3, \cdots, n_{(r-1)}),$$

where the summation extends over all $n_i[i=1, 2, 3, \cdots, (r-1)]$ which satisfy (2).

Application to "birthday problems." In applying this result to the "generalized birthday problems" posed earlier, one assumes $M=365$ possible birthdays in one year and

$$n_1 = \text{number of single birthdays}$$
$$n_2 = \text{number of double birthdays}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$n_{r-1} = \text{number of } (r-1)\text{-tuple birthdays.}$$

The solution to Problem 1 is determined by computing (5) for given n and r. To obtain a solution to Problem 2, one must choose an initial value for n and compute (5). If the probability is not equal to 1/2, then a second value of n is chosen to cause the probability in (5) to be nearer 1/2. This iterative process is continued until two consecutive positive integers are found such that the probability in (5) brackets 1/2. The larger of these two values is the number n such that the probability is greater than or equal to 1/2 that at least r people have the same birthday.

The successive use of (5) in determining the positive integer n satisfying Problem 2 has been carried out with the aid of an IBM 7090 computer. The results are tabulated in Table 1 below for $r=2, 3, 4$.

The author did not carry out the computation for $r=5$ since the machine time for one computation of (5) for a given n was estimated at two hours.

Some simplification of the program for computing (5) could be accomplished by identifying and disregarding negligible terms. It is questionable whether such considerations would appreciably reduce the computing time for $r>4$.

TABLE 1

r	2		3		4	
n	22	23	87	88	186	187
Prob.	.4758	.5074	.4998	.5114	.4965	.5033

Reference

1. Wm. Feller, An introduction to probability theory and its applications. Vol. I, 2nd ed. Wiley, New York, 1958.

ON THE INVERSE OF THE OPERATOR $\mathcal{A}(\cdot)=\mathbf{A}(\cdot)+(\cdot)\mathbf{B}$

J. P. JACOB AND E. POLAK, University of California, Berkeley

Introduction. In the course of studying (in [1]) the stability of certain classes of nonlinear systems, it is frequently necessary to invert the Lyapunov operator $\mathcal{L}$ which maps the space β of all $n\times n$ matrices into itself according to the rule:

$$\mathcal{L}(\mathbf{X}) = \mathbf{A}'\mathbf{X} + \mathbf{X}\mathbf{A}, \quad \text{for all } \mathbf{X} \in \beta,$$

where $\mathbf{A}$ is a $n\times n$ constant matrix and $\mathbf{A}'$ its transpose.

A more general form of this operator, $\mathcal{A}$, maps the space β into itself according to the rule:

$$\mathcal{A}(\mathbf{X}) = \mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{B}, \quad \text{for all } \mathbf{X} \in \beta,$$

where $\mathbf{A}$ and $\mathbf{B}$ are $n\times n$ matrices.

Although the necessary and sufficient conditions for the operator $\mathcal{A}$ to have an inverse are well known ([2], [3], [4], [5]) the authors feel that our alternative proof of these is somewhat simpler and may have pedagogical advantages.

First we recognize that the space β of all $n\times n$ constant matrices is a linear vector space. We now proceed to prove the following:

THEOREM. *Let $\mathcal{A}:\beta\to\beta$ be an operator defined by $\mathcal{A}(\mathbf{X})=\mathbf{A}\mathbf{X}+\mathbf{X}\mathbf{B}$, for all $\mathbf{X}\in\beta$, where $\mathbf{A}$, $\mathbf{B}$ are $n\times n$ constant matrices. Let λ_i: $i=1, 2, \cdots, m\leqq n$ be the eigenvalues of $\mathbf{A}$ and let μ_j: $j=1, 2, \cdots, l\leqq n$ be the eigenvalues of $\mathbf{B}$. Then $(\lambda_i+\mu_j)$, i, j, as above, is an eigenvalue of $\mathcal{A}$ and, conversely, any eigenvalue of $\mathcal{A}$, η_k: $k=1, 2, \cdots, p\leqq n^2$, can be represented as*

$$\eta_k = \mu_j + \lambda_i, \quad \textit{for some } i, j. \tag{1}$$

Proof. The operator $\mathcal{A}$, defined above, is a linear bounded operator. Linearity is obvious. Boundedness is trivial because β is finite dimensional. Note that $\mathcal{A}$ can be thought of as a linear operator of a "vector" $\mathbf{X}$ in an n^2 dimensional Euclidean space.

I. $(\lambda_i+\mu_j)$ is an eigenvalue of $\mathcal{A}$, where λ_i is any eigenvalue of $\mathbf{A}$, and μ_j is any eigenvalue of $\mathbf{B}$.

Let $\mathbf{x}_i$ and $\mathbf{y}_j$ be vectors such that

$$\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i, \tag{2}$$

$$\mathbf{y}_j'\mathbf{B} = \mu_j\mathbf{y}_j'. \tag{3}$$

Now, making use of (2) and (3), we obtain for the dyad $\mathbf{x}_i\times\mathbf{y}_j$,

$$\begin{aligned}\mathcal{A}(\mathbf{x}_i \times \mathbf{y}_j) &= \mathbf{A}\mathbf{x}_i \times \mathbf{y}_j + \mathbf{x}_i \times \mathbf{y}_j\mathbf{B}\\ &= \lambda_i\mathbf{x}_i \times \mathbf{y}_j + \mu_j\mathbf{x}_i \times \mathbf{y}_j\\ &= (\lambda_i + \mu_j)(\mathbf{x}_i \times \mathbf{y}_j).\end{aligned}$$

Hence $(\lambda_i+\mu_j)$ is an eigenvalue of $\mathcal{A}$ for all $i=1, 2, \cdots, m\leqq n, j=1, 2, \cdots, l \leqq n$.

II. All eigenvalues of $\mathcal{A}$ are of the form $(\lambda_i+\mu_j)$. Suppose η_k is any eigenvalue of $\mathcal{A}$, i.e., there exists a matrix $\mathbf{X}_k\neq\mathbf{0}$ such that

$$\mathcal{A}(\mathbf{X}_k) = \eta_k\mathbf{X}_k, \quad \text{i.e.,} \quad \mathbf{A}\mathbf{X}_k + \mathbf{X}_k\mathbf{B} = \eta_k\mathbf{X}_k.$$

Hence,

$$(\eta_k\mathbf{I} - \mathbf{A})\,\mathbf{X}_k = \mathbf{X}_k\mathbf{B}. \tag{4}$$

We now show that the matrices $(\eta_k\mathbf{I}-\mathbf{A})$ and $\mathbf{B}$ have at least one eigenvalue in common.

We prove this by contradiction. Suppose $(\eta_k\mathbf{I}-\mathbf{A})$ and $\mathbf{B}$ have no eigenvalue in common and let $g(\cdot)$ be the minimal polynomial of $(\eta_k\mathbf{I}-\mathbf{A})$. Then

$$g(\eta_k\mathbf{I} - \mathbf{A}) = 0 \quad \text{and} \quad g(\mathbf{B}) \neq \mathbf{0}. \tag{5}$$

Furthermore, $|g(\mathbf{B})|\neq 0$. This follows from the fact that

$$|g(\mathbf{B})| = \prod_{i=1}^{q} |(\mathbf{B} - \xi_i\mathbf{I})^{s_i}|,$$

where ξ_i, $i=1, 2, \cdots, q\leqq n$, are the eigenvalues of the matrix $(\eta_k\mathbf{I}-\mathbf{A})$ and s_i is the multiplicity of ξ_i as a root of the minimal polynomial $g(\cdot)$.

On the other hand, from (4), we have:

$$\begin{aligned}(\eta_k\mathbf{I} - \mathbf{A})^2\mathbf{X}_k &= (\eta_k\mathbf{I} - \mathbf{A})\mathbf{X}_k\mathbf{B} = \mathbf{X}_k\mathbf{B}^2\\ &\vdots\\ (\eta_k\mathbf{I} - \mathbf{A})^n\mathbf{X}_k &= \mathbf{X}_k\mathbf{B}^n.\end{aligned}$$

Consequently,

$$g(\eta_k\mathbf{I} - \mathbf{A})\cdot\mathbf{X}_k = \mathbf{X}_k\cdot g(\mathbf{B}), \tag{6}$$

$$\mathbf{0} = \mathbf{X}_k\cdot g(\mathbf{B}), \qquad \mathbf{X}_k = 0, \tag{7}$$

since we have proved $g(\mathbf{B})$ to be nonsingular. But this is a contradiction since $\mathbf{X}_k\neq 0$.

Therefore, since $(\eta_k\mathbf{I}-\mathbf{A})$ and $\mathbf{B}$ have at least one common characteristic value, we have, for some $i=1, 2, \cdots, m\leqq n$ and some $j=1, 2, \cdots, l\leqq n$

$$\eta_k - \lambda_i = \mu_j \quad \text{or,} \quad \eta_k = \lambda_i + \mu_j.$$

This completes the proof of the theorem.

COROLLARY 1. *The operator $\mathcal{A}$, defined in the theorem above, is nonsingular if and only if $(\lambda_i+\mu_j)\neq 0$ for all i, j.*

Proof. This is an immediate consequence of above theorem and definition of a linear nonsingular operator on a finite dimensional vector space.

COROLLARY 2. *The matrix equation* $\mathbf{AX}+\mathbf{XB}=\mathbf{C}$, *where* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *are constant* $n\times n$ *matrices, has a solution* $\mathbf{X}$ *for all* $\mathbf{C}$ *if and only if* $\lambda_i+\mu_j\neq 0$, $i=1, 2, \cdots, m \leqq n$, $j=1, 2, \cdots, l\leqq n$, *where* λ_i *is an eigenvalue of* $\mathbf{A}$ *and* μ_j *is an eigenvalue of* $\mathbf{B}$.

Proof. This is an immediate consequence of the Corollary 1, because the solution $\mathbf{X}$ of (7) is simply $\mathbf{X}=\mathfrak{A}^{-1}(\mathbf{C})$.

REMARK. A method of solving for $\mathbf{X}$ equations of the form $AX+XA'=C$ is presented by W. Givens [5]. Both R. Bellman [2] and W. Givens [5] obtain necessary and sufficient conditions for the solubility of this equation by identifying the Liapunov operator $\mathcal{L}(\cdot)$ with the operator $(A\otimes I+I\otimes B)(\cdot)$, where $\otimes$ denotes the tensor product of matrices.

Acknowledgement. This research was conducted at the Electronics Research Laboratory, University of California, Berkeley, and was supported in part by the Joint Services Electronics Program (Departments of the Air Force, Army, and Navy) under Air Force Office of Scientific Research Grant AF-AFOSR-139-64 and in part by the National Aeronautics and Space Administration Grant NsG-354 (S-1).

References

1. M. A. Aizerman and F. R. Gantmacher, Absolute Stability of Regulator Systems, C. II, Holden-Day, San Francisco, 1964.
2. R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.
3. W. Hahn, Eine Bemerkung zur zweiten Methode von Lyapunov, Math. Nachr., 14 Band., Heft 4/6, (1956) 349–354.
4. C. C. MacDuffee, The Theory of Matrices, Chelsea, New York, 1946, Chapter VII.
5. W. Givens, Elementary divisors and some properties of the Lyapunov mapping $\mathbf{X}\to\mathbf{AX}+\mathbf{XA}'$, Argonne Nat. Lab. Report ANL-6456, Nov. 1961.

ABSOLUTELY INDEPENDENT AXIOMS FOR THE DERIVED SET OPERATOR

SHAIR AHMAD, Western Reserve University

In a note [1], Harvey presented a set of axioms for the derived set operator, and stated that such an axiom system did not seem to be in print. It turns out, however, that many notions have long been discussed and analyzed in terms of the derived set operator (see [2] and [3]). For example, Kowalsky [3] introduced an equivalent set of axioms which may be stated as

I. $(A\cup B)^\alpha = A^\alpha \cup B^\alpha$, II. $A^{\alpha\alpha} \subset A\cup A^\alpha$,

III. $\varnothing^\alpha = \varnothing$, IV. *For all* $x\in X$, $x\notin \{x\}^\alpha$;

where A and B denote subsets of X.

It may be of interest to note that under a slight modification the above axiom system becomes "absolutely independent" in a sense indicated by Harary [4]. We observe that axiom III follows from I and IV, since $x\notin(\varnothing\cup\{x\})^\alpha =\varnothing^\alpha\cup\{x\}^\alpha$ implies that $x\notin\varnothing^\alpha$ for all $x\in X$. Furthermore, if $A=B$ then I is logically true. If $A=\varnothing$, then II follows from III which was proved from I and IV. Finally, II holds for any A such that $A'\subset\{x\}$ for some $x\in X$, where A'

denotes the complement of A in X. For, if $x \in A^\alpha$, then $A \cup A^\alpha = X$. If $x \notin A^\alpha$, then $A^\alpha \subset A$ which means that $A^{\alpha\alpha} \subset A^\alpha \subset A \cup A^\alpha$, since $(A^\alpha \cup A)^\alpha = A^{\alpha\alpha} \cup A^\alpha$.

In view of the above discussion, it is evident that the axioms

(1) $x \notin \{x\}^\alpha$ for all $x \in X$,

(2) $(A \cup B)^\alpha = A^\alpha \cup B^\alpha$ whenever $A \neq B$,

(3) $A^{\alpha\alpha} \subset A \cup A^\alpha$ for all A, such that $A \neq \emptyset$ and $A' \not\subset \{x\}$ for any $x \in X$

are equivalent to I, II, III and IV. Recalling that the "never holding" denials of (1), (2) and (3) in the sense used by Harary are

($\bar{1}$) There is at least one element $x \in X$, and $x \in \{x\}^\alpha$, for all $x \in X$,

($\bar{2}$) There are sets A and B contained in X such that $A \neq B$, and $(A \cup B)^\alpha \neq A^\alpha \cup B^\alpha$ whenever $A \neq B$,

($\bar{3}$) There is a set A contained in X such that $A \neq \emptyset$ and $A' \not\subset \{x\}$ for any $x \in X$; further $A^{\alpha\alpha} \not\subset A \cup A^\alpha$ whenever $A \neq \emptyset$ and $A' \not\subset \{x\}$ for any $x \in X$,

the following examples show that (1), (2) and (3) are "absolutely independent" with respect to the set $X = \{a, b, c\}$.

MAPPING	AXIOMS
$A^\alpha = \emptyset$	1 2 3
$A^\alpha = A$	$\bar{1}$ 2 3
$A^\alpha = A'$	1 $\bar{2}$ 3
$A = \emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, X$ $A^\alpha = \emptyset, \{b\}, \{c\}, \{a\}, \{b, c\}, \{a, b\}, \{a, c\}, X$	1 2 $\bar{3}$
$A = \emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, X$ $A^\alpha = X, \{c\}, \{a\}, \{b\}, \{b\}, \{a\}, \{c\}, \emptyset$	1 $\bar{2}$ $\bar{3}$
$\emptyset^\alpha = \emptyset$, $\{a\}^\alpha = \{a, b\}$, $\{b\}^\alpha = \{b, c\}$, $\{c\}^\alpha = \{a, c\}$, $A^\alpha = X$ for all other $A \subset X$	$\bar{1}$ 2 $\bar{3}$
$A^\alpha = A$ if A has exactly one element, otherwise $A^\alpha = A'$	$\bar{1}$ $\bar{2}$ 3
$\{a\}^\alpha = \{a, b\}$, $\{b\}^\alpha = \{b, c\}$, $\{c\}^\alpha = \{a, c\}$, $A^\alpha = A'$ for all other $A \subset X$	$\bar{1}$ $\bar{2}$ $\bar{3}$

It should be pointed out that while we thought it might be of interest to show that (1), (2) and (3) satisfy Harary's definition of an "absolutely independent" axiom system, we are not suggesting the justification of Harary's definition. There are difficulties involved in Harary's definition. The referee of this paper points out one such difficulty, namely that if we apply Harary's "never holding" denial to the two equivalent statements

(i) $(A \neq \emptyset$ and $A' \not\subset \{x\}) \to A^{\alpha\alpha} \subset A \cup A^\alpha$,

(ii) $A \neq \emptyset \to (A' \not\subset \{x\} \to A^{\alpha\alpha} \subset A \cup A^\alpha)$,

we get the two nonequivalent statements

($\bar{\text{i}}$) $(A \neq \emptyset$ and $A' \not\subset \{x\}) \to A^{\alpha\alpha} \not\subset A \cup A^\alpha$,

($\bar{\text{ii}}$) $A \neq \emptyset \to (A' \not\subset \{x\}$ and $A^{\alpha\alpha} \not\subset A \cup A^\alpha)$.

References

1. F. R. Harvey, The derived set operator, this MONTHLY, 70 (1963) 1085–1086.
2. Waclaw Sierpinski, Introduction to General Topology, Toronto, 1934.
3. H. J. Kowalsky, Topologische Räume, Birkhäuser Verlag, Basel und Stuttgart, 1961, p. 53.
4. Frank Harary, A very independent axiom system, this MONTHLY, 68 (1961) 159–162.

A NECESSARY AND SUFFICIENT CONDITION FOR CONVERGENCE OF REAL-VALUED FUNCTIONS

GIOVANNI VIDOSSICH, Avenza, Italy

In this paper, we give a necessary and sufficient condition for convergence of a real-valued function from which the classical convergence criteria can be easily deduced. Our notation and terminology are based on [1, Chap. ii], but we shall write:

$\mathbf{R}$ for the real line
$\overline{\mathbf{R}}$ for the extended real line
$]a, b[$ for the set $\{x \in \overline{\mathbf{R}} : a < x < b\}$, $a, b \in \overline{\mathbf{R}}$
f for the directed function $(f, \mathcal{A})$ when it is not important to know $\mathcal{A}$.

THEOREM 1. *Let f be a directed function in $\overline{\mathbf{R}}$. The conditions*
(a) *f converges to a point of $\overline{\mathbf{R}}$,*
(b) *there is at most one $x \in \mathbf{R}$ such that f is not ultimately $>$ nor $< x$,*
are equivalent to each other.

Proof. (a)$\Rightarrow$(b). This follows immediately from the definitions.

(b)$\Rightarrow$(a). Define $A = \{a \in \mathbf{R} : f$ *is ultimately* $> a\}$ and $B = \{b \in \mathbf{R} : f$ *is ultimately* $< b\}$.

CASE 1: $A = \emptyset$. Then $B = \mathbf{R}$. So f converges to $-\infty$.

CASE 2: $B = \emptyset$. Then $A = \mathbf{R}$. So f converges to $+\infty$.

CASE 3: A and $B \neq \emptyset$. Then, since $A \cap B = \emptyset$, $\bigvee A \leqq \bigwedge B$. If $\bigvee A < \bigwedge B$, for all $x \in]\bigvee A, \bigwedge B[$, f is not ultimately $>$ nor $< x$. This contradicts (b), so $\bigvee A = \bigwedge B$. Clearly, f converges to $\bigvee A = \bigwedge B$.

Theorem 1 is true even if in (b) we replace $>$ and $<$ by $\geqq$ and $\leqq$. The following corollaries constitute the well-known convergence criteria for real-valued functions.

COROLLARY 1. *If f is a directed function in $\mathbf{R}$, the conditions*
(a) *f converges to a point of $\mathbf{R}$,*
(b) *for every positive real number e, ultimately $|f(x') - f(x'')| < e$,*
are equivalent to each other.

Proof. (a)$\Rightarrow$(b). This follows from the definitions.

(b)$\Rightarrow$(a). Since (b) holds, f is ultimately bounded and so cannot converge to

$-\infty$ or $+\infty$. Hence, if (a) fails, f cannot converge to a point of $\overline{\mathbf{R}}$. Therefore by Theorem 1 there are distinct a and b in $\mathbf{R}$ such that f is frequently $\leqq$ and $\geqq$ both. But then (b) fails for $e=|a-b|$, a contradiction.

COROLLARY 2. *If* $(f, \mathcal{E})$ *is a directed function in* $\overline{\mathbf{R}}$ *and* $x\in\overline{\mathbf{R}}$, *the conditions*
(a) f *converges to* x,
(b) x *is the only cluster point of* f, *are equivalent to each other.*

Proof. (a)$\Rightarrow$(b). This follows immediately from the definitions.
(b)$\Rightarrow$(a). Since (b) holds, f cannot converge to a point of $\overline{\mathbf{R}}$ distinct from x. Hence if (a) fails, it follows from Theorem 1 that there are a and b in $\mathbf{R}$ such that $a<b$ and f is frequently $\leqq$ and $\geqq$ both. So

$$\mathcal{A} = \{f(E) \cap [-\infty, a]: E \in \mathcal{E}\} \quad \text{and} \quad \mathcal{B} = \{f(E) \cap [b, +\infty]: E \in \mathcal{E}\}$$

are directions. So, $\overline{\mathbf{R}}$ being compact, there are z and $y\in\overline{\mathbf{R}}$ such that $z\in\bigcap_{A\in\mathcal{A}} A$ and $y\in\bigcap_{B\in\mathcal{B}} B$. It is clear that $z\neq y$ and z, $y\in\bigcap_{E\in\mathcal{E}} f(E)$. Hence z and y are two different cluster points of f, i.e. (b) fails.

COROLLARY 3. *If* $(f, \mathcal{E})$ *is a directed function in* $\overline{\mathbf{R}}$, *the statements*
(a) f *converges to a point of* $\overline{\mathbf{R}}$,
(b) f *is order-convergent in* $\overline{\mathbf{R}}$, *are equivalent to each other.*

Proof. Define $x=\bigvee\{\bigwedge f(E): E\in\mathcal{E}\}$ and $y=\bigwedge\{\bigvee f(E): E\in\mathcal{E}\}$.
(a)$\Rightarrow$(b). (a) imples (b) of Theorem 1 and, since $x\leqq y$, this implies $x=y$.
(b)$\Rightarrow$(a). By definition of infimum and supremum, f is ultimately $\geqq(\leqq)$ all lower bounds (upper bounds) of $[x, y]$. So, since $x=y$, (b) of Theorem 1 holds, and hence (a) follows.

Reference

1. E. J. McShane and T. A. Botts, Real Analysis, Van Nostrand, Princeton, 1959.

A NOTE ON THE SCHUR-ZASSENHAUS THEOREM

WILLIAM LEAHEY, University of Illinois

Let G be a finite group of order $s\cdot t$ and suppose that G has a normal subgroup of order s. Then the Schur-Zassenhaus Theorem states that if s and t are relatively prime then G contains a subgroup of order t. If s and t are not relatively prime then of course G need not contain a subgroup of order t. However, if $s\cdot t$ is of the form $p\cdot q$ where p and q are primes or if $s\cdot t$ is of the form p^k with p prime, then the Sylow theorems imply that G always contains a subgroup of order t regardless of whether s and t are relatively prime. This is again the case if $s\cdot t$ is of the form $p^2\cdot q$, with p and q distinct primes, as can be seen by considering the various types of groups of order $p^2\cdot q$. (See, for example, [**1**, section 59].) Thus, if one wants an example of a group G of order $s\cdot t$ which contains a normal subgroup of order s but which does not contain a subgroup of order t, then the

smallest possible value for $s \cdot t$ is 24. The object of this note is to give a simple example of such a group of order 24.

Let G be the special linear group of 2×2 matrices with coefficients in GF(3), the finite field of three elements, i.e. G is the group of all 2×2 matrices A with coefficients in GF(3) which satisfy det $A=1$. Then order $(G)=24=2 \cdot 12$. The center Z of G, is a normal subgroup of order 2. It will be shown that G contains no subgroup of order 12.

First, by the Sylow theorems the number n of subgroups of order 3 is equivalent to 1 modulo 3 and is also a divisor of 24. Therefore either $n=1$ or $n=4$. But G contains at least three elements of order 3. For example, let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Then A, A^2 and B are three such elements. Therefore $n=4$.

Now suppose that G contains a subgroup H of order 12. Then H is normal in G and since G/H is of order 2, H must contain all elements of order 3. In particular, H contains A and B hence the element

$$C = (A \cdot B)^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

which is in Z. By taking the product of C by each of the eight elements of order 3 one has that H contains eight distinct elements of order 6. But then H contains eight elements of order 6 and eight elements of order 3 implying that order $(H) \geqq 16$, which contradicts the assumption that order $(H)=12$. Therefore G contains no subgroup of order 12.

Reference

1. W. Burnside, Theory of Groups of Finite Order, Cambridge University Press, 1911.

A SIMPLIFIED PROOF OF GAUSS-MARKOV THEOREM WHEN THE REGRESSION MATRIX IS OF LESS THAN FULL RANK

K. K. RAO, University of Minnesota

Introduction. Various proofs of the Gauss-Markov Theorem are available [**4, 5**] when the regression matrix is of less than full rank. All of them are rather lengthy and tedious. The simplest proof for the case of full rank is due to Plackett [**3**]. The same proof is extended here to the less than full rank case using the notions of generalized inverse of a matrix. The same technique has been used by Chipman [**1**] to develop the criterion of best linear minimum bias.

It can be shown [**2**] that for any $N \times K$ matrix X there exists a unique $K \times N$ matrix $X^\dagger$ satisfying

(i) $XX^\dagger X = X$ (ii) $X^\dagger XX^\dagger = X^\dagger$ (iii) $(XX^\dagger)' = XX^\dagger$ (iv) $(X^\dagger X') = X^\dagger X$

The matrix $X^\dagger$ is known as the generalized inverse of X. Consider an independent linear regression model: $y=X\beta+u$ with $E(u)=0$, $E(uu')=\sigma^2 I$ and X a given $N\times K$ matrix of arbitrary rank. Let $\psi\beta$ be an estimable (linear) function.

THEOREM (GAUSS-MARKOV). *The best linear unbiased estimator of $\psi\beta$ is given by $\psi X^\dagger y$.*

Proof. Let Ay be an unbiased estimator of $\psi\beta$, which implies $AX=\psi$. Consider the identity:

$$AA' - \psi X^\dagger X^{\dagger\prime}\psi' = (A-\psi X^\dagger)(A-\psi X^\dagger)' + (AX^{\dagger\prime} - \psi X^\dagger X^{\dagger\prime})\psi' + [(AX^{\dagger\prime} - \psi X^\dagger X^{\dagger\prime})\psi']'.$$

Using properties (ii) and (iii) we get

$$AX^{\dagger\prime} = AX^{\dagger\prime}X'X^{\dagger\prime} = AXX^\dagger X^{\dagger\prime} = \psi X^\dagger X^{\dagger\prime}.$$

Hence

$$AX^{\dagger\prime} - \psi X^\dagger X^{\dagger\prime} = 0 \quad \text{and} \quad AA' - \psi X^\dagger X^{\dagger\prime}\psi' = (A-\psi A^\dagger)(A-\psi X^\dagger)'.$$

Since $(A-\psi X^\dagger)(A-\psi X^\dagger)'$ is nonnegative definite we have the result: every diagonal of $AA' \geqq$ corresponding diagonal of $\psi X^\dagger X^{\dagger\prime}\psi'$. Q.E.D. The same proof goes through for the Aitken model also.

The author wishes to thank Professor John S. Chipman for his advice in the preparation of this note.

References

1. John S. Chipman, On least squares with insufficient observations, J. Amer. Statist. Assoc., 59 (1964) 1078–1111.

2. R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc., 51 (1955) 406–413.

3. R. L. Plackett, A historical note on the method of least squares, Biometrika, 36 (1949) 458–460.

4. C. Radhakrishna Rao, Generalization of Markov's Theorem and Tests of Linear Hypotheses, Sankhyā, 7 (1945) Part I, 9–16.

5. Henry Scheffé, The Analysis of Variance. Wiley, New York, 1959.

A NOTE ON EMBEDDING COMMUTATIVE RINGS IN RINGS WITH UNITY

MICHAEL GEMIGNANI, University of Notre Dame and St. Mary's College

Although it is a well-known fact that any ring can be embedded as an ideal in a ring with unity, the purpose of this note is to introduce a device, which the author has not yet seen in print, by which we can produce for a "nice" commutative ring A a "natural" ring with unity in which A can be embedded as an ideal.

Let A be a nonzero commutative ring. Define $T(A)=\mathrm{Hom}_A(A, A)$, the ring of all A-endomorphisms of the additive group of A. Then $T(A)$ is a ring with unity. For each a in A, let $\bar{a}$ be the left multiplication determined by a. Then

$a \to \bar{a}$ defines a homomorphism of A onto a subring $\bar{A}$ of $T(A)$. It is easily verified that $\bar{A}$ is an ideal in $T(A)$. If A has unity 1, then $T(A) = \bar{A}$.

THEOREM. *If for each nonzero a in A, there is b in A such that $ab \neq 0$, then*
(a) *$a \to \bar{a}$ is an embedding;*
(b) *$T(A)$ is commutative;*
(c) *$T(A)$ is an integral domain iff A has no divisors of* 0; *and*
(d) *if $n = \text{char}A$, then $n = \text{char}T(A)$.*

Proof: (a) The hypothesis of this theorem is equivalent to assuming that the kernel of the homomorphism defined by $a \to \bar{a}$ is $\{0\}$.

(b) For any a, b in A and f, g in $T(A)$ we have $f \circ g(ab) = f(g(ab)) = f(g(a)b) = [f(g(a))]b = [f \circ g(a)]b$ and $f \circ g(ab) = f(g(ab)) = f(ag(b)) = f(a)g(b) = g(b)f(a) = g(bf(a)) = g(f(a)b) = [g(f(a))]b = [g \circ f(a)]b$, hence $[(f \circ g - g \circ f)(a)]b = 0$. Since for any given a, we have b such that $[(f \circ g - g \circ f)(a)]b \neq 0$ if $(f \circ g - g \circ f)(a) \neq 0$, it must be that $f \circ g = g \circ f$.

(c) Clearly if A has divisors of zero, then $T(A)$ could not be an integral domain. Suppose that A has no divisors of zero. Then $T(A)$ is commutative by (*b*). Suppose there are nonzero f, g in $T(A)$ such that $f \circ g = \bar{0}$. Then we can find a, b in A such that $f(b) \neq 0$ and $g(a) \neq 0$. Therefore $f \circ g(ab) = [f \circ g(a)]b = 0b = 0 = g(a)f(b)$, a contradiction.

(d) Since A can be embedded in $T(A)$, $\text{char}A \mid \text{char}T(A)$. But $nf(a) = f(na) = f(0)$ for all f in $T(A)$ and all a in A, hence $\text{char}T(A) \leqq \text{char}A$, and the result follows at once.

A CHARACTERIZATION OF THE SUPREMUM

A. C. WILLIAMS, Socony Mobil Oil Company, Inc. Princeton, N. J.

Let X be a subset of some universal set U, and let f be a function with domain F in U and with range in the extended real numbers. Then sup $\{f(x) \mid x \in F \cap X\}$ may be considered as a function S with range in the extended real numbers and with domain pairs (X, f) consisting of a set $X \subseteq U$ and a function f from $F \subseteq U$ to the extended reals. When $F \cap X$ is empty we assign to S the value $-\infty$. We may call S the sup function of X and f.

The sup function satisfies the following three conditions:

1. If $F \cap X$ is not empty, and if $f(x) = c$ for all $x \in F \cap X$, then $S(X, f) = c$. If $F \cap X$ is empty, then $S(X, f) = -\infty$.
2. If $Y \subseteq X$, then $S(Y, f) \leqq S(X, f)$.
3. If $g \geqq f$ on $F \cap X$, then $S(X, g) \geqq S(X, f)$.

THEOREM. *The sup function is the only function from pairs (X, f) to the extended reals for which conditions* 1, 2, *and* 3 *are satisfied, i.e. conditions* 1, 2, *and* 3 *characterize the sup function.*

Proof. Let S denote any function from pairs (X, f) to the extended reals, for which conditions 1, 2, and 3 are satisfied. Consider any (X, f) and let g be the function on $F \cap X$ whose value is everywhere equal to sup $\{f(x) \mid x \in F \cap X\}$. Then, by conditions 3 and 1:

$$S(X, f) \leqq S(X, g) = \sup\{f(x) \mid x \in F \cap X\}.$$

Moreover, from 2 and 1: $S(X, f) \geqq S(\{x\}, f) = f(x)$, each $x \in F \cap X$. Therefore, $S(X, f) \geqq \sup \{f(x) \mid x \in F \cap X\}$.

ANOTHER MIXED NONGROUP

J. W. Wyman, Pasadena College

A counter-example showing that a set which is closed under an associative binary operation and which has a left identity and right inverses for all the elements need not be a group was given by Colonel Johnson Jr., in this Monthly, 71(1964)785. Another counter-example is the following:

Let A be a proper subset, with more than one element, of a set B and $G = \{f: B \to A \mid f(x) = x \text{ for each } x \in A\}$.

(i) G is closed under the operation of composition of functions.

(ii) This operation is associative.

(iii) Let $x_1 \in A$. The function $g: B \to A$ defined by $g(x) = x$ if $x \in A$, $g(x) = x_1$ if $x \in B - A$, belongs to G and serves as a left identity since, for each $f \in G$, $gf = f$.

(iv) If $f \in G$, g serves as a right inverse for f with respect to g, since $fg = g$.

The element g is not a right identity since A has more than one element and G does not form a group under the given operation.

Editor's Note: Another reader, Mr. Gary A. Yoshioka, points out that the table

$$\begin{array}{c|cc} & f & g \\ \hline f & f & f \\ g & g & g \end{array}$$

defines a "mixed nongroup." If $X = \{x, y\}$, then the mappings

$$f: \begin{matrix} x \to x \\ y \to x \end{matrix} \quad \text{and} \quad g: \begin{matrix} x \to y \\ y \to y \end{matrix}$$

satisfy this table.

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and
JOHN A. BROWN, University of Delaware

*All material for this department should be sent to John R. Mayor
1515 Massachusetts Avenue, N. W., Washington, D. C. 20005.*

A DILEMMA IN MATHEMATICAL EDUCATION

T. L. SAATY, Conference Board of the Mathematical Sciences

Traditionally no field of knowledge equals mathematics in the fascination and excitement which it evokes in its practitioners. Once under its magical spell an individual cannot leave it without strong feeling of desertion—perhaps even a sense of diminishing stature. This rank of mathematics is accomplished through the work of people. They are the masters and innovators, individuals who have put themselves in harmony with the field by creative work in it. Some are discoverers, some are educators, and even some are critics and connoisseurs.

Since both the status and development of mathematics require an *esprit de corps* and a strong camp of followers who are mindful of each other in a constructive spirit, it is valuable to take a look at an aspect of mathematical education which might improve this solidarity both in method and in end result. Most mathematicians are potentially fine in human values and equally willing to give of themselves to see others obtain fulfillment by being mathematically creative. Occasionally one is inclined to feel that he must make progress all by himself. For some educators there is perhaps a passing feeling of discouragement with the available raw material.

A distinguished colleague, in an offguarded witty utterance confided this discouragement by saying: "Universities are a wonderful place to be, but the worst thing about them is that they have students." I seriously doubt that he genuinely felt this way.

In science and even more in mathematics, the rigors and high standards required of students take so much of the energy of their mentors to enforce, that little time is left for consideration of personal growth. A clear perspective of the total person is needed.

I have known many a would-be great scholar and creator of ideas who has fallen short of the attainment of his goals, not because of lack of great intelligence or talent or even self discipline, but mainly because of the absence of needed personal warmth to temper an imagined or existing nonconcern or perhaps harshness in the environment. This is particularly characteristic of some of the larger departments. The United States has always been distinguished by the generous spirit of its people towards their fellowmen. It is surprising that this spirit has not been adequately and maturely interwoven in the fabric of higher education in science. It is usually taken for granted that for producing worthy

scholars, a good mind is most of what is needed and personal development is secondary. However, there are enough examples of brilliant individuals who failed because of lack of personal development that we need to re-examine this philosophy in order to decrease what I consider a substantial waste of talent, and a nontrivial loss.

Many sensitive and talented individuals learn from early childhood to live in a world of their own intellectual making. In spite of depth of a model of the natural world, the component having to do with human relations may be a naive one. Consequently, the task of growing up to assume responsibility and being effective in a real world of people is painful and for some the desire for philosophical and personal accommodation occurs at the time that a student is receiving his graduate education. At that time he is approaching a peak of his awareness of the surrounding world. Thus he faces two problems—one revolves about excellence in demonstrating his creative mental acumen and the other in adapting his "world picture" to a framework in which he can exist harmoniously. These are both strong forces. The intelligent man in this category is sure of his ability to cope with academia, but is more apprehensive about how his picture fits the world. Failure at school may be due to apathy resulting from the fact that he attaches greater significance and exerts greater effort in sorting out his relations in the world than in the fulfillment of his academic obligations to a point where his interests in the pursuit of his creative talents may cease, a result of confusion of objectives.

Assuming that this diagnosis is sufficiently valid, let us examine briefly a possible external remedy. First, it would be helpful if his teachers, who exert perhaps the strongest influence on his total development, were intellectual in outlook and had a broader awareness of life than the distinguished colleague of our introductory remark.

Of course, there are many individuals who are successful scientists who may never have faced the problem of resolving a complicated mental picture of the world. By example, they serve to integrate and inspire a student who is anxious for interpretations and encouraging suggestions.

There are also those teachers who, almost like intellectual giants, have found an accommodation for both their academic excellence and the world about them. They are the ones who must be singled out by various imaginative means to help the student through this difficult period. Every school and every department must have examples of them. To the type of student under discussion they are living examples of what he might make of himself. Without them his education may make him technically proficient, but still he must look elsewhere for a gift of human values and for a better realization of himself. A feeling of urgency in this self realization and absence of the catalytic encouragement of experienced teachers may precipitate a crisis before his academic career is finished. To be an effective teacher one must care about the student.

Some individuals who are absorbed in research are by nature such that they do not enjoy being frequently bothered but may allow questions along their

line of specialization. They may be ones who have learned when to exercise mental curiosity about the world and when to block it in pursuit of academic excellence. For some of their students whose curiosity may be strong, blocking would not work. Therefore, such teachers would not be in a favorable position to understand the problem and render help at a period of concentration and these periods are known to last a long time.

Perhaps another area of improvement rests with a better exploration of a student's talents. Some students arrive at a creative apex of their field by different roads, but others are lost because at their school ordinarily only one road is available, a certain agreed upon road which simplifies the life of the educator, but not of the student. To be sure brilliance and depth do not lack with a large number of those who do not make the mark, but they just don't have the knack of walking in the footsteps of their elders or in absorbing themselves along traditional lines.

I know a number of examples of very capable students who left school or were asked to leave because they could not *adapt* themselves to the regime or learn at the hands of personally colorless "masters" or immature specialists. In many cases, perhaps especially in the smaller schools, a course seemed to mean that the expert must again go through the chore of regurgitating old stuff or unloading on the blackboard, with brilliance and with the touch of a Chef of French Cuisine, that which they have memorized from many previous presentations. The style is so perfect that interruptions with questions which might beset the student are regarded as boorish and lacking in the manner of high scholarship. Thus, the student assumes the role of a copier from the board to the paper and gains little insight from the teacher's experience. His education is mostly achieved at home. In fact since the teacher copies his material from books and published papers, it would be better training for the student to look things up for himself and spend the time in class in discussions which might illuminate points of special difficulty. He may even find out that there are things in the course which the teacher himself does not know, thus giving him the idea that one need not be perfect to be creative.

It is well known that fashions in science are changeable and that in any case at the current rapid pace of growth in 50 or 100 years what is most likely to survive is the spirit scientists convey as carriers of the torch. This indicates that rather than spending much time in perfecting the old, one perhaps should acquire and communicate the zest of the new.

The foregoing discussion indicates that to produce effective teachers, more than specialization in a field is required. Knowledge and practice in teaching methods, courage and constructiveness in human relations are essential ingredients. They provide the fine distinction between assembly line production of talent and a respect for the higher and deeper value of young people.

SEARCHING FOR MATHEMATICAL TALENT IN WISCONSIN, II

J. R. SMART, University of Wisconsin

For the second year the Department of Mathematics of the University of Wisconsin has conducted a program of searching for high school students with mathematical talent (see [1]). The project was again supported financially by the National Science Foundation. The purpose of the project is to locate and encourage highly talented students. The program is as follows. This year we sent out five sets of problems with five problems per set. The first set was mailed in September and the last in mid-March. The students had approximately four weeks to work on each set. In the Appendix we list the problem sets along with the number of responses, and in parenthesis after each problem is the number of correct solutions. The problem sets were mailed to all 559 high schools in the state, both public and parochial. Approximately 31,000 copies of each set were distributed. After the problems were graded we prepared and mailed in equal numbers a sheet containing solutions and a list of solvers using the format of the problem sections of this MONTHLY. There were 626 individuals who responded to one or more problems. The breakdown of this number by grade in school is as follows: 12—235, 11—186, 10—124, 9-68, 8—12, 7—1.

In May the top forty-five problem solvers and their teachers were invited to a one day program on the Madison campus. All of the students as well as twenty-one teachers attended. We arranged for the students to be broken up into groups of four each and they met with a faculty member for forty-five minutes to one hour. This occurred in both the morning and afternoon. In addition, in the morning Professor I. J. Schoenberg gave a one hour lecture entitled "On the Kakeya Problem," and in the afternoon the students and teachers toured the computing facilities of the University. A luncheon was served at the home of Professor L. C. Young and there was a dinner at the Wisconsin Memorial Union. At the close of the dinner, for further encouragement, each student was presented with paperback copies of *How to Solve It* by George Polyá, *Mathematical Recreations and Essays* by W. W. Rouse Ball, and *Enrichment Mathematics for High School* published by the National Council of Teachers of Mathematics. In addition, the top six students Jeffrey Kuester, Gregory Markowski, Daniel Shapiro, Michael Schweitzer, Richard Snelling and John Wasserstrass were presented with either *What is Mathematics* by R. Courant and H. Robbins or *Geometry and the Imagination* by D. Hilbert and S. Cohn-Vossen. The students who were invited to this program solved six or more problems; the top six students solved fourteen to eighteen problems!

The Mathematical Talent Search has been quite well received by students and teachers throughout the state. We get many favorable comments from each group. There are some additional aspects of this program which deserve mentioning. We know, from students' and teachers' comments, that many additional students attempted the problems and that quite a few of these obtained solutions which they did not bother to send in. Thus we are reaching and en-

couraging more students than the numbers quoted above indicate. Another fact which may have great value in the future is that lines of communication have opened between high school teachers and the faculty of the Mathematics Department. This has already proved valuable in obtaining financial assistance for deserving students. Finally I should like to mention that twenty-five of the forty-five top problem solvers were seniors and seventeen of these have decided to attend the University of Wisconsin.

This year the program was run by Professor Michael Bleicher, Professor Donald Crowe and the author. The program will continue next year with Professor Crowe as chairman. If we can be of assistance to others in setting up a similar program in other states, we would be most happy to do so.

Reference

1. M. N. Bleicher, Searching for mathematical talent in Wisconsin, this MONTHLY, 72 (1965) 412–416.

APPENDIX

The problem sets are listed below exactly as they were sent out. We realize now that some of the problems could be more effectively worded.

Problem Set 1—406 Responses

1. In a certain town, the blocks are rectangular, with the streets (of zero width) running E–W, the avenues N–S. A man wishes to go from one corner to another m blocks East and n blocks North. The shortest path can be achieved in many ways. How many? (22)

2. An army captain wants to station an observer equally distant from two points and a road. Can it be done? If so, how many points are there and can one locate these points on a map using straight edge and compass? Give the construction, if possible. In other words, how many points are there in the Euclidean plane which are equidistant from two given points and a given line? Find them with straight edge and compass, if possible. (18)

3. For which values of $n \geqq 0$ and $k \geqq 0$ is

$$(k+1)^n + (k+2)^n + (k+3)^n + (k+4)^n + (k+5)^n$$

divisible by 5? (19)

4. Represent the number 1 as the sum of reciprocals of finitely many *distinct* integers larger than or equal to 2. Can this be done in more than one way? If so, how many? (22)

5. Six points in space are such that no three are in a line and no four are in a plane. The fifteen line segments joining them in pairs are drawn and then painted, some segments red, some blue. Prove that some triangle formed by the segments has all its sides the same color. (10).

Problem Set 2—235 Responses

1. Show that if a, b, c, are three integers which satisfy $a+b\sqrt{2}+c\sqrt{3}=0$, then $a=b=c=0$. (71)

2. Give a method for the construction (with straight-edge and compass) of a triangle with given angles α and β, $\alpha+\beta<180°$, and given perimeter s. (17)

DEFINITION. *A prime number is defined to be a positive integer larger than* 1 *which is only divisible by itself and* 1.

The Fundamental Theorem of Arithmetic: *Every integer larger than* 1 *can be factored as a product of powers of prime numbers and the factorization is unique (except for the ordering of the factors). Thus any integer n larger than* 1 *can be factored as* $p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}$ *where* $p_1, p_2, \cdots p_k$ *are the distinct prime divisors of n and* $a_1, a_2, \cdots, a_k$ *are positive integers.*

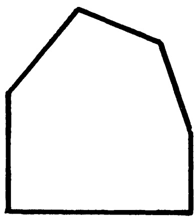
3. Show that for $n\geqq 5$, 24 divides the product

$$n^2\left(1-\frac{1}{p_1^2}\right)\left(1-\frac{1}{p_2^2}\right)\cdots\left(1-\frac{1}{p_k^2}\right),$$

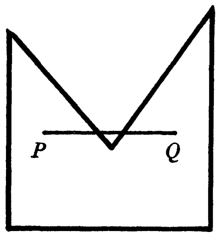
where $p_1, p_2, \cdots, p_k$ are the distinct prime divisors of n. (12)

4. Let $a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_n$ be $2n$ positive real numbers. Show that either $a_1/b_1+a_2/b_2+\cdots+a_n/b_n\geqq n$ or $b_1/a_1+b_2/a_2+\cdots+b_n/a_n\geqq n$. (12)

DEFINITION: *A polygon (or figure) is called convex if for each two points on or inside the polygon (or figure) the line segment joining these two points is on or inside the polygon (or figure).*

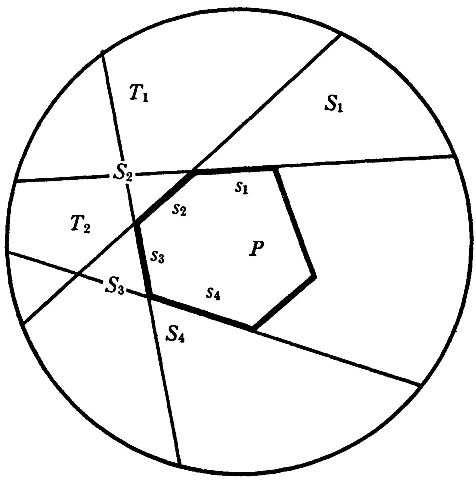


convex



not convex

5. A convex polygon P of area $\bar{P}$ is contained in the interior of a circle C of area $\bar{C}$. Let the sides of P be numbered counterclockwise: $s_1, s_2, \cdots, s_n$. Each side is extended

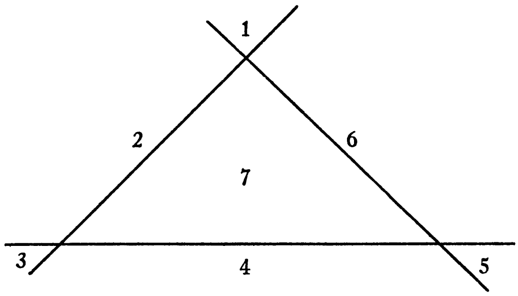


in both directions so that it divides the circle into two segments. Let S_i denote the segment determined by s_i which does not contain P. Let $\overline{S}_i$ denote the area of S_i and $\overline{T}_i$ the area common to S_i and S_{i+1} $(i=1,2, \cdots, n, S_{n+1}=S_1)$.
Show that

$$\overline{P} = \overline{C} - \overline{S}_1 - \overline{S}_2 - \cdots - \overline{S}_n + \overline{T}_1 + \overline{T}_2 + \cdots + \overline{T}_n. \quad (30)$$

Problem Set 3—200 Responses

1. The figure shows three lines dividing the plane into seven regions. Find the maximal number of regions into which the plane can be divided by n lines. (13)



2. (a) Let $a_1, a_2, \cdots, a_n$ be n positive numbers. Show that if $a_1+a_2+ \cdots +a_n \leqq 1$, then $1/a_1+1/a_2+ \cdots +1/a_n \geqq n^2$.

(b) Now suppose $a_1, \cdots, a_n$ are arbitrary positive numbers (with unrestricted sum). Show that

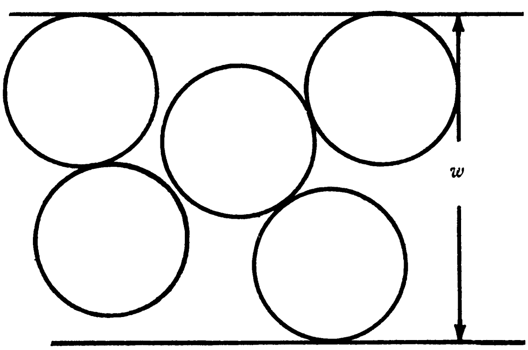
$$(a_1 + a_2 + \cdots + a_n)(1/a_1 + 1/a_2 + \cdots + 1/a_n) \geqq n^2. \quad (6)$$

3. Let $p_1, p_2, \cdots, p_n$ be n points located in the plane. Show that the shortest broken line connecting the points does not cross itself. (22)

4. Let $P(x)=a_nx^n+a_{n-1}x^{n-1}+ \cdots +a_1x+a_0$ be a polynomial of degree $n \geqq 1$ with integer coefficients. Show that there are infinitely many positive integers m such that

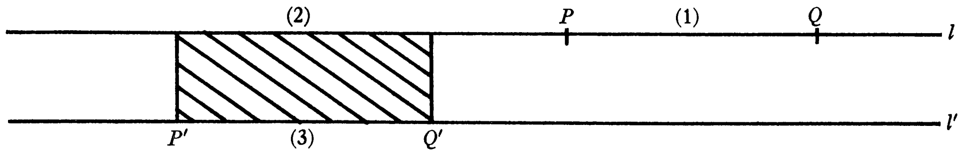
$$P(m) = a_nm^n + a_{n-1}m^{n-1} + \cdots + a_1m + a_0 \text{ is not prime.} \quad (3)$$

5. We consider a set of non-overlapping circles of radius 1 lying in a parallel strip of width w. We say that the circles form a k-cloud ($k=1, 2, 3, \cdots$) if any line intersecting the strip intersects at least k circles. Show that for a 2-cloud $w \geqq \sqrt{3}+2$. (Two circles are said to overlap if they have a common inner point.) (40)



Problem Set 4—122 Responses

1. Suppose l and l' are two parallel lines one inch apart, and $\overline{PQ}$ and $\overline{P'Q'}$ are segments each two inches long, on l and l' respectively. Then clearly $\overline{PQ}$ can be moved in the



plane of l and l' so as to coincide with $\overline{P'Q'}$ and in the process sweeps out an area of exactly two square inches. Prove that if l and l' are 100 miles apart it is still possible to move $\overline{PQ}$ to coincide with $\overline{P'Q'}$ while sweeping out a total area of no more than two square inches. (65)

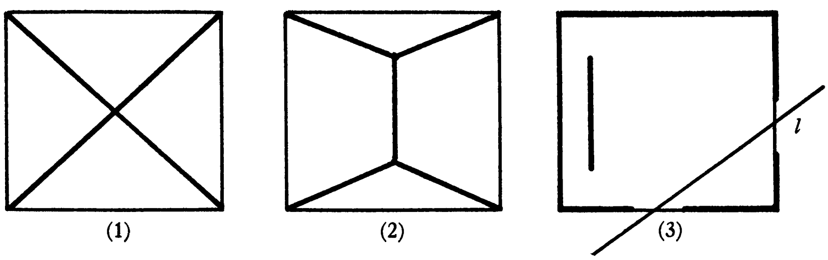
2. Show that N can be taken so large that $1+1/2+1/3+1/4+ \cdots +1/N$ is larger than 100. (37)

3. Let R be a point of a polygon p such that any segment joining R with a point of p lies completely in p. Then p is said to be *visible* from R. Prove that the set of all points from which p is visible is convex (a set of points in the plane is convex if given any two points of the set then the line segment joining the points lies in the set—see problem set 2). (19)

DEFINITION. *Two integers m and n are said to be relatively prime if they have no common factors (i.e.,* 1 *is the only positive integer dividing both m and n). The integers in a collection of integers are said to be pairwise relatively prime if every pair of integers from the collection are relatively prime.*

4. Let a collection $a_1, a_2, \cdots, a_n, \cdots$ of integers be defined successively by $a_{n+1}=a_n^2-a_n+1$ and $a_1=2$. The first few are $a_1=2$, $a_2=3$, $a_3=7$, $a_4=43$, $a_5=1807, \cdots$. Show that the integers $a_1, a_2, a_3, \cdots$ are pairwise relatively prime. (14)

5. A collection of line segments contained in a square of side length one inch is said to be an opaque set if every straight line which crosses the square cuts one of the line segments. Examples (1) and (2) are opaque sets while (3) is not. Think of the line segments



as painted walls in a house otherwise made of glass. For an opaque set a flashlight beam cannot pass through the house. An opaque set need not consist of connected line segments (as in (1) and (2)). The length of the opaque set (1) is $2\sqrt{2}$ (≈ 2.828) inches, and the length of opaque set (2) is $1+\sqrt{3}$(≈ 2.732) inches. Find an opaque set which is less than $1+\sqrt{3}$ inches. (58)

Problem Set 5—47 Responses

1. Consider the collection of integers $a_1, a_2, \cdots$ defined by $a_{n+1}=a_n^2-a_n+1$, $a_1=2$. Show that by taking N very large $|1/a_1+1/a_2+\cdots+1/a_N-1|<1/10^{10}$. (8)

2. Let P be an arbitrary polygon in the plane (the sides do not intersect). Describe how to construct with compass and straight-edge a square having the same area. (5)

3. Let m and n be relatively prime integers (see problem set 4) with $0<m/n<1$. Find a method for testing when a given m/n can be expressed as $1/x+1/y$ where x and y are integers. (12)

4. Finitely many pennies are placed on a flat surface, no two over-lapping. Prove or disprove: No matter how this is done it is always possible to paint each penny with one of three colors in such a way that no two pennies having the same color touch each other. (0)

An n by n square matrix H is an array of real numbers of the form

$$\begin{bmatrix} a_{11} & a_{12} \cdots a_{1\,n-1} & a_{1n} \\ a_{21} & a_{22} \cdots a_{2\,n-1} & a_{2n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{n1} & a_{n2} \cdots a_{n\,n-1} & a_{nn} \end{bmatrix}$$

The numbers a_{ij} are called the *entries* in the matrix. The first subscript tells which row the number is in, and the second subscript, the column. For example, a_{36} is in the third row and sixth column. Two rows, say $(a_{j1}, a_{j2}, \cdots, a_{j\,n-1}, a_{jn})$ and $(a_{k1}, a_{k2}, \cdots, a_{k\,n-1}, a_{kn})$ $j\neq k$, are said to be *orthogonal* if

$$a_{j1}a_{k1}+a_{j2}a_{k2}+\cdots+a_{j\,n-1}a_{k\,n-1}+a_{jn}a_{kn}=0.$$

5. A matrix H all of whose entries are $+1$ or -1 and whose rows are orthogonal is called a Hadamard matrix after the famous French mathematician Jacques Hadamard, 1865–1963. An example is

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

Show that there are infinitely many Hadamard matrices by constructing one of order 2^n for each n. (We remark that it is an unsolved problem of whether there is a Hadamard matrix of order 116 or not.) (8)

TEACHING AND RESEARCH CONFLICT

Federal research programs and the nation's goals for higher education are in "increasing conflict," according to a congressional subcommittee report (the House Subcommittee for Research and Technical Programs) on the effects of government research and development programs on higher education. The subcommittee headed by Rep. Henry S. Reuss (D., Wis.) also criticized the

concentration of federal research funds in a relatively few major institutions to the detriment of smaller colleges, and to the overemphasis on natural sciences as compared with humanities and the social sciences.

The conflict between research support and goals of higher education identified by the subcommittee fell roughly in three categories. These are: allocation of scarce manpower, conflict between present use of manpower for research and investment for future manpower needs as exemplified by teaching, and conflict within the colleges between the natural sciences, humanities, and social sciences.

One of the biggest needs identified by the committee is for more incentives for teaching. Several recommendations were made to try to meet this need. The committee called for science teaching fellowships which would be as attractive financially as research fellowships, for a Presidential award for outstanding undergraduate teachers, and for limiting requirements for research manpower over a short term as incentives are provided to increase the flow of qualified persons into teaching.

The report pointed out that colleges and universities bear a responsibility for maintaining a proper balance between teaching and research.

BRIEF COMMENT

Proceedings of the Preliminary Meeting on College Level Mathematics Education under the Auspices of the U.S.—Japan Program on Scientific Cooperation, Katada, 1964. Japan Society for the Promotion of Science, 1965. xii+96 pp.

This booklet is a report of a "lively meeting held under the auspices of the U.S.-Japan Program on Scientific Cooperation" from September 5 through September 7, 1964 in Japan. The meeting was attended by twelve Japanese and four American mathematicians and they "examined aspects of the college mathematics curriculum—What topics should there be, how taught, and in what order." A partial table of contents follows:

It is hoped that a number of the above reports will be reviewed in this column in future issues.

The Madison Project's Approach to a Theory of Instruction, ROBERT B. DAVIS, *J. Res. Sci. Teaching*, 2(1964) 146–162.

Apparently this is an attempt to develop a theory of instruction based upon considerable classroom success of the Madison Project's methods of mathematics instruction. There is no well developed theory, but the ideas are broadly clustered around Jean Piaget's notion that education is fundamentally a process in which the individual's misperceived and incorrect view of the real world goes through a successive sequence of refinements and improvements so that a more sophisticated view is achieved. The Madison Project has attempted to bring the student at the elementary level in contact with mathematics with as little mediation from the teacher as possible, and has chosen topics with the following aims in mind: the student must be ready for the ideas and must take an active role in developing them; the concepts must arise naturally from some problem solving situation, must be related to some fundamental mathematical ideas, and must lead to some significant patterns of generalization; the topics must be appropriate to the age of the child, and must appear, *in toto*, to the observer to amount to a significant experience. There is also considerable emphasis on a non-rigid, flexible classroom attitude on the part of the teacher and upon reinforcement that comes from within the student rather than from the teacher. There is much in the article that is thought provoking and interesting even though it may not bear directly upon the development of a theory of instruction.

Computer-Aided Instruction, JOHN A. SWETS and WALLACE FEURZEIG, *Science*, October 29, No. 3696, 150 (1965).

Although this article does not deal with the teaching of mathematical concepts, it is not hard to extrapolate the techniques used here to such usage. The basic teaching technique is the "Socratic System" in which the computer states a problem and thereafter the student and computer engage in a dialog while the student attempts to solve the problem. The dialog consists of questions by the computer at various times, answers to questions posed by the student on the part of the computer, and suitable comments by the computer. The student has a basic vocabulary consisting of a set of questions and statements which he may use in the dialog—the machine is programmed to respond in a suitable fashion depending upon the specific response on the part of the student plus what has gone before in the dialog.

The principal example used in this article is the teaching of medical diagnosis of a particular ailment. It is clear, however, that the technique is quite flexible and that it would adapt well to many rather complex concepts. "To play this game, the Socratic System has only to recognize logical conditions, such as sufficiency, necessity, redundancy, and consistency. This capability is not inconsequential—for example such a capability might be sufficient in an aid for teaching fundamentals of geometry, or qualitative inorganic chemistry—but it can be realized in a less complex system."

The Madison Project—A Brief Introduction to Materials and Activities, ROBERT B. DAVIS, The Madison Project, Webster College, St. Louis, Missouri, 1965, 31 pp.

The Madison Project which began about eight years ago seeks to improve the school mathematics curriculum through broadening it to include more advanced concepts, introducing a more creative flavor, providing more variety, and encouraging a more active student participation. There is considerable emphasis on informal instruction—particularly the early introduction of topics informally before they are discussed formally. Notations of the UICSM materials are used, and stress is placed on the point that students seem to learn the notion of variable more easily if some non-alphabetic symbol, such as a box, is used rather than the conventional "x" or "y". The project does not exist primarily to provide and produce teaching materials, but rather to provide services and to pursue various activities. The project has worked with various school systems in both large and small communities in carrying out its philosophy of mathematical education and has also engaged in teacher education activities. The pamphlet contains a detailed bibliography of publications available from the project, as well as cognate publications. The pamphlet is available from The Madison Project, Webster College, St. Louis, Missouri 63119.

Mathematics and the High School Curriculum, F. V. POHLE, Department of Graduate Mathematics, Adelphi University, 1965, 41 pp.

This report, written by the chairman of the department of graduate mathematics at Adelphi University, appears to be largely an apologia for and an expansion of the views of Professor Morris Kline as expressed in the Alumni News of New York University, October, 1961. The recommendations of the report are these: better training for teachers of mathematics, increased cooperation between college and high school teachers, greater unification in the teaching of science and mathematics, caution in even the suggestion for curriculum changes, and lack of need for any massive changes in subject matter in the high school curriculum. The beginning of the report makes interesting reading in that it details some of the events in the controversy over changes in the high school curriculum—there are some cogent, as well as amusing, remarks by such people as Kline, Moise, Stoker, and others. Unfortunately the report, particularly the last section on recommendations, is marred by serious typographical errors, woolly thinking, and poor English usage.

Copies of the report are obtainable without charge from Professor Pohle, Adelphi University, Garden City, N. Y. 11530.

Editorial Note. "Brief comment" was offered (this MONTHLY 72 (1965) 901) on an article by R. W. Hamming in SCIENCE (148 (1965) 473–475) entitled "Numerical Analysis vs. Mathematics." Two critical letters and a rejoinder appear in SCIENCE, 149 (1965) 243–245.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE, Bloomfield College

COLLABORATING EDITORS: J. BARLAZ, Rutgers—The State University; L. CARLITZ, Duke University; H. S. M. COXETER, University of Toronto; H. EVES, University of Maine; A. E. LIVINGSTON, University of Alberta; and A. WILANSKY, Lehigh University.

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, Bloomfield College, Bloomfield, N. J. 07003. Proposers of Problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to A. E. Livingston, University of Alberta, Edmonton, Alberta, Canada. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before August 31, 1966.

E 1875. *Proposed by Samuel Zaidman, University of Montreal*

Let $\phi\in C^\infty[0, 1]$ satisfy the following conditions:

$$(1) \qquad \phi \not\equiv 0, \qquad (2) \qquad \phi^{(k)}(0) = 0, \quad k = 0, 1, \cdots,$$

(3) The series $\sum a_k \phi^{(k)}(x)$ is uniformly convergent on $[0, 1]$ where $\{a_k\}$ is a given sequence of real numbers. Prove that $k!a_k \to 0$ as $k \to \infty$.

E 1876. *Proposed by K. K. Norton, University of Illinois*

Let $\phi(n)$ be Euler's totient function, let $d(n)$ be the number of divisors of n, and let $\sigma(n)$ be the sum of those divisors. Show that

$$(1) \qquad \liminf_{n\to+\infty} \frac{\phi(n)}{nd(n)} = \liminf_{n\to+\infty} \frac{\sigma(n)}{nd(n)} = 0,$$

$$(2) \qquad \limsup_{n\to+\infty} \frac{\phi(n)}{nd(n)} = \limsup_{n\to+\infty} \frac{\sigma(n)}{nd(n)} = \tfrac{1}{2}.$$

(Cf. E 1625 [1964, 683] and E 1749 [1966, 84].)

E 1877. *Proposed by Huseyin Demir, Middle East Technical University, Ankara*

Let $ABCDE$ be a convex pentagon inscribed in a unit circle with AE as diameter, and let $AB=a$, $BC=b$, $CD=c$, $DE=d$. Then prove that

$$a^2 + b^2 + c^2 + d^2 + abc + bcd < 4.$$

E 1878. *Proposed by Huseyin Demir, Middle East Technical University, Ankara*

Let $A_1A_2 \cdots A_n$ be a regular polygon inscribed in circle (O) of radius R. Denote the incenter of the triangle $A_{i-1}A_iA_{i+1}$ (indices mod n) by I_i, and that

of the triangle formed by A_iA_{i+2}, $A_{i+2}A_{i+1}$, $A_{i+1}A_{i+3}$ by J_i. Then show that the points $I_1, \cdots, I_n$ and $J_1, \cdots, J_n$ all lie on the same circle of radius $R' = R\cos(3\pi/2n)/\cos(\pi/2n)$.

E 1879. *Proposed by Agatha Himmelfarb, Fordham University*

A billiard ball is cued from a corner of an $n\times m$-foot billiard table at an angle of 45° (n and m are integers.) How many cushions will the ball strike before it again goes into a corner?

E 1880. *Proposed by Robert Schlesinger, Washington University, St. Louis, Mo.*

Show that $f(t)=(t-\sin t)(\pi-t-\sin t)$ is an increasing function of t for $0<t<\frac{1}{2}\pi$.

E 1881. *Proposed by Omar Khayyam, Jr., University of California, Berkeley*

Describe the locus of a point P in the Euclidean plane for which there is a function f with the property that the graph of a function g defined on the whole real line is symmetric with respect to P if and only if g commutes with f (i.e., $f(g(x))=g(f(x))$ for all real x.)

E 1882. *Proposed by W. C. Beckham, Kirtland Air Force Base, New Mexico*

Let I represent the $n\times n$ unit matrix. A is an $n\times n$ matrix which has one and only one nonzero element in each row and column, these nonzero elements being either 1 or -1. If $A^2=I$, determine how many different ways there are of writing A.

E 1883. *Proposed by A. M. Vaidya, Texas Technological College, Lubbock*

The Möbius function $\mu(n)$ provides a criterion for deciding whether or not a given integer is squarefree. Find a function (in terms of μ) which has different constant values for primes, for composite square-free integers, and for other integers.

E 1884. *Proposed by A. F. Beardon, University of Maryland*

Prove that the series

$$\sum_{n_1,\ldots,n_k=1}^{\infty}\frac{1}{(n_1^2+\cdots+n_k^2)^p}$$

converges if and only if $p>k/2$.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Consequence of Pell's Equation

E 1757 [1965, 182]. *Proposed by J. L. Brown, Jr., Ordnance Research Laboratory, State College, Pennsylvania*

Show that the equation $t''=2t'$ has infinitely many solutions in triangular numbers t' and t''.

I. *Solution by Lacy G. Jones, Ripley High School, Ripley, West Virginia.* This problem has been solved many times before (cf. Dickson's *History of Theory of Numbers*, Volume II, p. 182). Therefore, the following solution will no doubt be a repetition of what has already been done.

The equation $t'=2t''$ can be written $\{x(x+1)\}/2=y(y+1)$ which can be reduced by $x^2+(x+1)^2=z^2$, where $z=2y+1$. Now, if we can generate an infinite number of right triangles such that one of the legs is one unit greater than the other, we will have solved our problem. This may be accomplished by the transformation $X=2z+3x+1$, $Z=3z+4x+2$ which will yield an infinite number of right triangles with $X^2+(X+1)^2=Z^2$, where $X>x$ and $Z>z$.

II. *Solution by F. D. Parker, SUNY, Buffalo, New York.* The problem can be generalized: If α is a positive integer, then $t'=\alpha t''$ has infinitely many solutions in r-gonal numbers t' and t'' if α is not a perfect square or if $r=4$ and α is a perfect square.

The nth r-gonal number is $\{(r-2)n^2-(r-4)n\}/2$, so that we seek solutions for the Diophantine equation $(r-2)n^2-(r-4)n=\alpha(r-2)k^2-\alpha(r-2)k$. It can be verified that

$$n=\frac{r-4}{r-2}\,\frac{1\pm(x+4\alpha(r-2)y)}{2},$$

$$k=\frac{r-4}{r-2}\,\frac{1\pm(x+4(r-2)y)}{2},$$

provide a solution if x and y are integer solutions to $x^2-16\alpha(r-2)^2y^2=1$. This is a Pellian equation and has an infinite number of solutions. Since $(x-1)(x+1)=16\alpha(r-2)^2y^2$, x is odd and $x-1$ and $x+1$ have no common factor except 2, so that either $x-1$ or $x+1$ has a factor $r-2$. Since $r-2\neq 2$, both n and k are integers.

If $r=4$, there are an infinite number of solutions if α is a square integer.

The problem as originally stated is the case in which $r=3$, $\alpha=2$.

Also solved by A. N. Aheart, J. W. Baldwin, M. G. Beumer (Netherlands), M. T. L. Bizley (England), D. A. Blaeuer, Walter Bluger and W. Zayachkowski (jointly), E. O. Buchman, J. A. Burslem, Mannis Charosh, P. L. Chessin, M. R. Chowdhury (West Germany), D. I. A. Cohen, Monte Dernham, D. Z. Djokovic (Yugoslavia), R. B. Eggleton (Australia), G. E. Engebretsen, Michael Fredman, Michael Goldberg, S. H. Greene, L. S. Grinstein, Cornelius Groenewoud, G. E. Henry, Agatha Himmelfarb, J. A. H. Hunter, R. F. Jackson, Erwin Just, M. S. Kaplan, E. S. Langford, E. L. Magnuson, Andrzej Makowski (Poland), D. C. B. Marsh, Yehoshua Mayer ben-David, R. J. Oberg, C. B. A. Peck, Harsh Pittie, Judith Richman, Ron Rietz, P. A. Scheinok, C. D. Scudder III, Robert and Robin Sibson (jointly) (England), R. L. Syverson, G. C. Thompson, Guy Torchinelli, Simon Vatriquant (Belgium), Lenard Weinstein, R. A. Wiesen, Dale Woods, and the proposer.

A Maximal Convex Envelope

E 1758 [1965, 182]. *Proposed by Michael Goldberg, Washington, D. C.*

Arrange four radial lines of lengths $a\geqq b\geqq c\geqq d$ so that their convex envelope is maximized.

I. *Solution by D. C. B. Marsh, Colorado School of Mines.* Assuming that the problem means maximize the area of the convex quadrilateral with vertices at the endpoints of the radial lines, we consider an ordering s_1, s_2, s_3, s_4 of any four radial segments with angles of θ_j between s_j and $s_{j+1}(s_5=s_1)$ and find the area of the envelope to be

$$A = (s_1s_2 \sin\theta_1 + s_2s_3 \sin\theta_2 + s_3s_4 \sin\theta_3 + s_4s_1 \sin\theta_4)/2.$$

By the use of Lagrange multipliers, or other standard methods, the maximum A is found to occur when all $\theta_j=\pi/2$. $A=(s_1+s_3)(s_2+s_4)/2$. For the s_j a permutation of $a\geqq b\geqq c\geqq d$, we then determine the maximum area to exist for the arrangement a, b, d, c, (i.e., with a, d at 180°, b, c at 180°, and ad at right angles to bc).

II. *Solution by Richard A. Jacobson, South Dakota State University, Brookings, South Dakota.* We maximize the length of the convex envelope. Suppose four lines of length r_1, r_2, r_3, r_4 are arranged counter-clockwise and are separated by angles $\theta_1, \theta_2, \theta_3, \theta_4$. The envelope has length

$$\sum_{i=1}^{4} \sqrt{(r_i^2 + r_{i+1}^2 - 2r_ir_{i+1}\cos\theta_i)},$$

where $r_5=r_1$. This is certainly a maximum if $\cos\theta_i=0$. Thus we let the $\theta_i=\pi/2$. The problem is then reduced to finding which permutation produces a maximum length. There are essentially only three distinct permutations $a-b-c-d$, $a-b-d-c$, and $a-d-b-c$. The respective lengths are

$$(1)\quad L_1 = \sqrt{(a^2+b^2)} + \sqrt{(b^2+c^2)} + \sqrt{(c^2+d^2)} + \sqrt{(d^2+a^2)}$$

$$(2)\quad L_2 = \sqrt{(a^2+b^2)} + \sqrt{(b^2+d^2)} + \sqrt{(d^2+c^2)} + \sqrt{(c^2+a^2)}$$

$$(3)\quad L_3 = \sqrt{(a^2+d^2)} + \sqrt{(d^2+b^2)} + \sqrt{(b^2+c^2)} + \sqrt{(c^2+a^2)}.$$

Cancelling like terms, squaring, again cancelling, squaring, and utilizing the condition that $a\geqq b\geqq c\geqq d$, we find that $L_3\geqq L_1\geqq L_2$. Thus the envelope is maximized when the lines radiate from the origin and terminate at the points $(a, 0)$, $(0, d)$, $(-b, 0)$ $(0, -c)$.

Also solved by E. O. Buchman, D. I. A. Cohen, D. Z. Djokovic (Yugoslavia), P. K. Garlick, S. H. Greene, R. F. Jackson, E. S. Langford, Charles McCracken, C. B. A. Peck, Judith Richman, Robin Sibson (England), T. Teichmann, and the proposer.

Symmetrically Placed, Nonattacking Rooks

E 1759 [1965, 182]. *Proposed by M. J. L. Bizley, London, England*

In how many ways can n rooks be placed on an $n\times n$ chessboard so that no rook can take another and so that the n occupied squares are symmetrically placed relative to a given diagonal of the board?

Solution by C. B. A. Peck, State College, Pennsylvania. Let there be m rooks, with $0\leqq m\leqq n$. Given a diagonal, we choose $n-m$ of its squares to be un-

occupied and unattacked. This can be done in $\binom{n}{m}$ ways. We then choose $m-2k$ of the remaining m diagonal squares for occupation, in any one of $\binom{m}{2k}$ ways, where k is some integer with $0 \leqq 2k \leqq m$. Finally, we place the remaining $2k$ rooks in pairs symmetrical about the given diagonal so that each pair has two of the remaining $2k$ diagonal squares for its exclusive joint attack, which we can do in $(2k)!/k!2^k$ ways. Then all together there are, for m rooks on an $n \times n$ board,

$$\sum_{k=0}^{[m/2]} \binom{n}{m}\binom{m}{2k}\frac{(2k)!}{2^k k!}$$

ways, with the case of the problem being $m=n$.

Also solved by W. A. Al-Salam, Robert Bart, John Beidler, E. O. Buchman, D. I. A. Cohen, R. J. Cormier, R. B. Eggleton (Australia), G. E. Engebretsen, D. Z. Djokovic (Yugoslavia), D. M. Hancasky, R. F. Jackson, Agnis Kaugars, D. C. B. Marsh, H. F. Mattson and R. J. Turyn (jointly), W. E. Patten, Donald Quiring, Robin Sibson (England), Guv Torchinelli, Simon Vatriquant (Belgium), and the proposer.

A Convergent Sequence Arising from a Difference Equation

E 1760 [1965, 183]. *Proposed by I. I. Kolodner, University of New Mexico*

$\{a_n\}$ and $\{b_n\}$ are sequences of real numbers such that $\{b_n\}$ converges and $b_n = a_n - \alpha a_{n+1}$. Show that $\{a_n\}$ converges if $|\alpha| > 1$, or if $|\alpha| < 1$ and $\{a_n\}$ is bounded.

Solution by Sidney Spital, California State Polytechnical College, Kellogg, Pomona, California. We prove the more general

THEOREM. *If $\{a_n\}$ and $\{b_n\}$ are sequences of complex numbers such that $\{b_n\}$ converges and $b_n = a_n - \alpha a_{n+1}$ with α a fixed complex number, then $\{a_n\}$ is convergent if $|\alpha| > 1$ or if $|\alpha| < 1$ and* $\lim_{n\to\infty} a_n\alpha^n = 0$.

Proof. The conclusion is clear when $\alpha = 0$. For $\alpha \neq 0$, we may view $a_n - a_{n+1} = b_n$ as a first order, linear, nonhomogeneous difference equation in the unknown sequence $\{a_n\}$. Its solution is known to be

$$a_n = -\sum_{k=-1}^{n-1} b_k\left(\frac{1}{\alpha}\right)^{n-k} = -\frac{1}{\alpha^n}\sum_{k=-1}^{n-1} b_k\alpha^k, \tag{1}$$

where for notational consistency we have labelled the arbitrary starting value $a_0 = -b_{-1}/\alpha$.

If $|\alpha| > 1$ we use the first part of relation (1) which shows that $\{a_n\}$ are partial sums of a power series in $1/\alpha$ with coefficients $-b_k$, $k = \cdots, 1, 0, -1$. Since $\{b_k\}$ is convergent, these coefficients are bounded. The convergence of the power series and $\{a_n\}$ then follows from $|1/\alpha| < 1$.

If $|\alpha| < 1$ we use the second part of (1):

$$-a_n\alpha^n = \sum_{k=-1}^{n-1} b_k\alpha^k.$$

Since $a_n\alpha^n$ tends to zero the partial sums on the right also approach zero. They may therefore be replaced by the negatives of their corresponding remainders

$$-a_n\alpha^n = -\sum_{k=n}^{\infty} b_k\alpha^k \quad \text{or} \quad a_n = \sum_{k=0}^{\infty} b_{n+k}\,\alpha^k.$$

Writing $\lim_{n\to\infty} b_n = b$, we estimate the difference

$$\left| a_n - \frac{b}{1-\alpha} \right| = \left| \sum_{k=0}^{\infty} (b_{n+k} - b)\alpha^k \right| \leqq \frac{1}{1 - |\alpha|} \sup_{k\geqq n} |b_k - b|,$$

which therefore must become arbitrarily small for increasing n. Hence $\{a_n\}$ converges to $b/(1-\alpha)$.

For $\alpha=1$, $a_n = -\sum_{k=-1}^{n-1} b_k$, and the example $b_{-1}=b_0=-1$ and $b_n=(-1)^k/2^k$ for $2^k \leqq n < 2^{k+1}$ $(k=0, 1, 2, \cdots)$ shows that $\{a_n\}$ can be bounded without being convergent. Similarly, $\alpha=-1$ and $b_n\equiv 1$ leads to $a_n=[1+(-1)^n]/2$.

Also solved by E. O. Buchman, Jim Campbell, Robert Cohen, D. F. Dawson, D. Z. Djokovic (Yugoslavia), W. O. Egerland, J. A. Goldstein, D. M. Hancasky, D. A. Hejhal, Stephen Hoffman, R. F. Jackson, Ben Klein, E. S. Langford, D. C. B. Marsh, Yehoshua Mayer ben-David, C. B. A. Peck, L. J. Pratte, J. G. Rau, Robin Sibson (England), Richard Sinkhorn, Al Somayajulu, C. D. Soong (Malaysia), G. P. Speck, Guy Torchinelli, Albert Wehrly, A. Wilansky, and the proposer.

Editorial Note. For $|\alpha|>1$, the first of equations (1) in Spital's solution defines a sequence-to-sequence linear summability scheme with matrix (a_{nk}), where $a_{nk}=1/\alpha^{n+1-k}$ for $k\leqq n$ and $a_{nk}=0$ for $k>n$, and the Silverman-Toeplitz conditions are easily seen to be satisfied. Several solvers based their solutions on this observation.

Wehrly found a similar problem in G. H. Hardy, *Pure Mathematics*, p. 168, and Wilansky calls attention to his book *Functional Analysis*, Blaisdell (1964), paragraph 1.2, problems 20, 21 22 and paragraph 14.3, problem 16.

Another Vandermonde Determinant

E 1761 [1965, 183]. *Proposed by John Burke, University of Vermont*

Let A be an $n\times n$ matrix with $a_{ij}=j^{i-1}$. Prove that

$$\det A = (n-1)(n-2)^2 \cdots 2^{n-2}.$$

Solution by Gary L. Musser, Student, SUNY, Buffalo, New York. This is the Vandermonde determinant $V(x_1, x_2, \cdots, x_n)$ with $x_j=j$ and, hence, has the value

$$\prod_{i>j} (i-j) = \prod_{k=2}^{n-1} k^{n-k}.$$

Also solved by J. C. Abad, H. D. Abramson, A. N. Aheart, R. G. Albert, W. A. Al-Salam, J. H. Avila, Jr., K. F. Bailie, Robert Bart, John Beidler, M. A. Bershad, M. G. Beumer (Netherlands), Marjorie Bicknell, M. T. L. Bizley (England), D. A. Blaeuer, J. A. Burslem, F. A. Butter, Jr., M. R. Chowdhury (Germany), D. I. A. Cohen, R. J. Cormier, H. J. de St. Germain, D. Z. Djokovic (Yugoslavia), W. G. Dotson, Jr., H. E. Dunsmore, W. O. Egerland, R. B. Eggleton (Australia), G. E. Engebretsen, L. O. Ferguson, Philip Fung, P. K. Garlick, Mrs. A. C. Garstang,

Michael Goldberg, Robert Gouw, S. H. Greene, L. S. Grinstein, D. M. Hancasky, J. Z. Hearon, W. E. Hoff, Stephen Hoffman, R. F. Jackson, P. G. Kirmser, M. S. Klamkin, H. E. Lahmann (Germany), E. S. Langford, J. C. Lazzard and Jack Barone (jointly), E. L. Magnuson, Andrzej Makowski (Poland), D. C. B. Marsh, Gus Mavrigian, Yehoshua Mayer ben-David, C. F. McLaren, M. R. Meck, A. P. Mella, M. G. Murdeshwar, F. D. Parker, C. B. A. Peck, J. M. Perry, Simeon Reich (Israel), Henry Ricardo, D. P. Roselle, J. P. Ruebsamen, P. A. Scheinok, Robin Sibson (England), R. Sivaramakrishnan (India), F. C. Smith, J. M. Smith, Sidney Spital, R. L. Syverson, Guy Torchinelli, B. R. Toskey, Simon Vatriquant (Belgium), Albert Wehrly, Lenard Weinstein, Charles Wells, R. A. Wiesen, H. H. Wong, Theodore Zerger, and the proposer.

$\zeta^2(2)/\zeta(4)$

E 1762 [1965, 183]. *Proposed by Underwood Dudley, University of Michigan*

Evaluate

$$\sum_{m=1,\,(m,n)=1}^{\infty} \sum_{n=1}^{\infty} \frac{1}{m^2n^2}.$$

Solution by N. J. Fine, Pennsylvania State University, University Park, Pennsylvania. For $k_1, \cdots, k_r > 1$, define

$$S(k_1, \cdots, k_r) = \sum_{(n_1, \cdots, n_r)=1} n_1^{-k_1} n_2^{-k_2} \cdots n_r^{-k_r}.$$

Then

$$\begin{aligned}
\zeta(k_1 + \cdots + k_r)S(k_1, \cdots, k_r) &= \sum_{a=1}^{\infty} a^{-(k_1+\cdots+k_r)} \sum_{(n_1, \cdots, n_r)=1} n_1^{-k_1} \cdots n_r^{-k_r} \\
&= \sum_{a=1}^{\infty} \sum_{(n_1, \cdots, n_r)=1} (an_1)^{-k_1} \cdots (an_r)^{-k_r} \\
&= \sum_{a=1}^{\infty} \sum_{(m_1, \cdots, m_r)=a} m_1^{-k_1} \cdots m_r^{-k_r} = \sum_{m_1, \cdots, m_r=1}^{\infty} m_1^{-k_1} \cdots m_r^{-k_r} \\
&= \zeta(k_1) \cdots \zeta(k_r).
\end{aligned}$$

Hence

$$S(k_1, \cdots, k_r) = \frac{\zeta(k_1) \cdots \zeta(k_r)}{\zeta(k_1 + \cdots + k_r)}.$$

In the given problem, $k_1 = k_2 = 2$, so $S(2, 2) = \zeta^2(2)/\zeta(4)$. Since $\zeta(2) = \pi^2/6$, $\zeta(4) = \pi^4/90$, the required sum has the value $5/2$.

Also solved by M. G. Beumer (Netherlands), W. J. Blundon, D. I. A. Cohen, Louis Comtet (France), D. Z. Djokovic (Yugoslavia), Glenn E. Engebretsen, P. K. Garlick, S. H. Greene, M. G. Greening (Australia), Roger Grimson, Peter Hagis, Jr. and Leon Steinberg (jointly), R. F. Jackson, M. S. Klamkin, E. S. Langford, Andrzej Makowski (Poland), D. C. B. Marsh, J. J. Martinez, Yehoshua Mayer ben-David, C. B. A. Peck, Michael Rosen, P. A. Scheinok, F. G. Schmitt, Robin

Sibson (England), Alan Slomson (England), C. P. Soong (Malaysia), Guy Torchinelli, A. M. Vaidya, and the proposer.

Beumer, Makowski, and Slomson point out that this problem appears as Number 9 in Nieuw Archief voor Wiskunde, XII (1964) 68.

Fine's argument of course proves the

THEOREM. *If* $f(kn_1, kn_2, \cdots, kn_r)=g(k)f(n_1, n_2, \cdots, n_r)$ *when* $(n_1, n_2, \cdots, n_r)=1$ *and the series*

$$S = \sum_{n_1, n_2, \cdots, n_r \geqq 1} f(n_1, n_2, \cdots, n_r)$$

is absolutely convergent, then

$$\sum_{(n_1, n_2, \cdots, n_r)=1} f(n_1, n_2, \cdots n_r) = S \sum_{k=1}^{\infty} g(k).$$

Random Shuffling by Computers

E 1763 [1965, 183]. *Proposed by E. O. Thorp, New Mexico State University*

A computer programmer attempts to shuffle N cards by the following process. Let i be a random variable integer uniformly distributed over the integers $1 \leqq i \leqq N$. Generate an i. Interchange the first card and ith card. Then generate another i. Interchange the second card with the ith card. Continue until the Nth card has been interchanged with the ith card. For which $N \geqq 2$ is the shuffling random?

Solution by D. Z. Djokovic, University of Belgrade, Yugoslavia. The final arrangement of cards is uniquely determined by the sequence $(i_1, i_2, \cdots, i_N)$, where i_k is the kth generated i. All these N^N sequences have the same probabilities. If the shuffling is random the N^N sequences must split (according to the final arrangement of cards) into $N!$ classes containing the same number of sequences. Hence, N^N must be divisible by $N!$. The only possible value is $N=2$. In this case the shuffling is random since the sequences (1, 2), (2, 1) give the original arrangement of cards and the sequences (1, 1), (2, 2) give the inverse arrangement.

Also solved by E. O. Buchman, T. J. Burke, G. E. Engebretsen, H. T. Gaines, R. F. Jackson, E. S. Langford, D. C. B. Marsh, J. M. Perry, Donald Quiring, Judith Richman, Daniel Ritchie, M. J. Sheridan, Robin Sibson (England), Mitchell Snyder, and the proposer.

A Consequence of E 1592

E 1764 [1965, 183]. *Proposed by Michael Gemignani, University of Notre Dame*

Let G be a group with identity 1, and A a subgroup of G such that $(G-A)\cup\{1\}$ is also a group. Prove that either $A=\{1\}$ or $A=G$.

I. *Solution by David M. Cohen, Student, Central High School, Philadelphia.* Let $B=(G-A)\cup\{1\}$ and note that A and B have only 1 in common. For $a\in A$ and $b\in B$, there is a c in G for which $ac=b$. If $c\in A$, then $b\in A$ since A is a group and, hence, $b=1$; in other words, $B=\{1\}$ and $A=G$. On the other hand, if $c\in B$, then c^{-1} and, hence, $a=bc^{-1}$ are in B, so that $a=1$ and $A=\{1\}$.

II. *Solution by Robert Gouw, Student, University of Hawaii.* According to Problem E 1592 [this MONTHLY 70 (1963) 568, and 71 (1964) 319], the union of two subgroups is a group if and only if one contains the other. Here, $A \subset (G-A) \cup \{1\}$ imples that $A = \{1\}$, while $(G-A) \cup \{1\} \subset A$ gives $G-A=\varnothing$.

III. *Solution by Jack C. Mettauer, West Virginia University, Morgantown, West Virginia.* A *groupoid* is a set G which is closed under an operation usually denoted by juxtaposition; it is *cancellative* if there is a cancellation law for the groupoid operation. The present problem is then a special case of the

THEOREM. *If $G=A\cup B$ is a cancellative groupoid with A a groupoid in which the equation $ax=a'$ has a unique solution x in A for a, $a' \in A$ and B is a groupoid in which the equation $yb=b'$ has a unique solution y in B for b, $b' \in B$, then $A \subset B$ or $B \subset A$.*

Proof: Suppose that $a \in A-B$ and $b \in B-A$. If $ab \in A$, there is an $a' \in A$ for which $aa'=ab$. But then $a'=b \in A$, a contradiction. Similarly, $ab \in B$ implies that $b'b=ab$ for some b' in B and, then, $b'=a \in B$, a contradiction. Thus, one of $A-B$ and $B-A$ is empty.

Also solved by J. C. Abad, R. G. Albert, J. S. Alin, D. J. Allen, W. A. Al-Salam, R. M. Aron and Richard Resch (jointly), D. R. Arterburn, T. A. Atchison, K. F. Bailie, Raymond Balbes, Drayton Barbare, C. W. Barnes, Cornelius Baytop and James Joseph (jointly), M. A. Bershad, Dennis Bertholf, Bruce Blum, W. E. Bodden, F. S. Brenneman, Robert Brook, J. L. Brown Jr., S. H. Brown, E. O. Buchman, Jim Campbell, P. L. Chabot, M. R. Chowdhury (West Germany), John Christopher, C. C. Clever, D. I. A. Cohen, Irma Cohen and Edward Cohen (jointly), Robert Cohen, Coker College Senior Math Seminar, R. J. Cormier, Peter Day, J. T. Dillon, Lyle Dixon, D. Z. Djokovic (Yugoslavia), W. G. Dotson, Jr., Naomi Dunner, R. B. Eggleton (Australia), G. E. Engebretsen, E. W. Ewing, J. R. Fall, W. F. Feeny, L. O. Ferguson, David Finkel, Philip Fung, Hyman Gabai, D. M. Giblin, Anton Glaser, K. R. Goodearl, Ed Gotway, Jr., M. G. Greening (Australia), Cornelius Groenewoud, J. D. Haggard and D. W. Hight (jointly), D. M. Hancasky, R. B. Hardin, Fr. Dunstan Hayden, D. A. Hejhal, G. A. Heuer, Agatha Himmelfarb, W. E. Hoff, Stephen Hoffman, G. T. Hogan, Edward Hook, R. J. Hursey, Jr., R. F. Jackson, W. D. Jackson, R. A. Jacobson, Thomas Jefferson, Jr., Erwin Just, Agnis Kaugars, B. G. Klein, Richard Kowalski, Kenneth Kramer, C. D. La Budde, H. E. Lahmann (Germany), E. S. Langford, Kenneth Lebensold, Steve Ligh, C. C. Lindner, Robert Madell, J. J. Martinez, D. C. B. Marsh, J. L. Matucha, C. J. Maxson, Yehoshua Mayer ben-David, C. F. McLaren, M. R. Meck, R. E. Mikhel, S. E. Minear, Sister Irene Morvan, J. G. Moser, M. G. Murdeshwar, J. B. Muskat, R. M. Nelson, R. J. Oberg, F. J. Papp, Jr., Robert Patenaude, C. B. A. Peck, Donald Perlis, Harsh Pittie, J. R. Porter, Donald Quiring, M. A. Rasheed (India), G. J. Rau, Fred Rosenblum, J. M. Rosenstein, Bernard Rosner, L. E. Rudsinski, P. J. Ryan, M. S. R. K. Sastry, Perry A. Scheinok, R. N. Schneider, Gerald Schrag, Hilbert Schultz, G. J. Schumm, L. L. Scott, M. J. Sheridan, D. R. Shoemaker, Robert Sibson (England), C. P. Singer, Indranand Sinha, Richard Sinkhorn, R. Sivaramakrishnan (India), F. E. Siwiec, D. A. Smith, A. L. Somayajulu, Robert Spira, R. E. Stockton, T. F. Sullivan, J. P. Tarwater, D. E. Tepper, Guy Torchinelli, B. R. Toskey, J. J. Uhl, A. M. Vaidya, William Vale, C. Van de Vyle (Belgium), A. G. Vassalotti, Nelda Viser, John Waddington, Daniel Warner, C. Watari (Japan), Albert Wehrly, Lenard Weinstein, Gary B. Weiss, Charles Wells, B. A. Welsh, J. C. Wenger, R. E. Whitney, D. A. Wick and S. R. Johnson (jointly), C. R. Williams, J. C. Williams, B. B. Winter, and the proposer.

Many of the above solutions were incomplete in that the solvers assumed G to be finite.

Schumm and Smith call attention to problem 30, p. 80, of I. N. Herstein, *Topics in Algebra*, Blaisdell (1964).

Problem 3739 Again

E 1765 [1965, 183]. *Proposed by Andy Vince, Stanford University*

Given a set of $n+1$ integers chosen from the set 1, 2, 3, $\cdots$, $2n$, it is always possible to choose two elements from the $n+1$ integers such that one divides the other.

Solution by John F. Dillon, Department of Defence, Fort Meade, Maryland. If A is any set of nonzero integers, let us call a subset S of A *admissible* if $x \nmid y$ and $y \nmid x$ whenever $x, y \in S$ and $x \neq y$.

THEOREM. *Let $T(a)$ denote the greatest odd integer divisor of the nonzero integer a. If A is a finite set of nonzero integers and $T(A)$ has ν distinct members, then each admissible subset of A has at most ν distinct members.*

Proof: Defining $a \sim b$ to mean that $T(a)=T(b)$, it is clear that $\sim$ is an equivalence relation on A, which partitions A into ν disjoint equivalence classes. Evidently, $x|y$ or $y|x$ if x and y are in one of these equivalence classes. Consequently, any admissible subset of A has at most one member in each equivalence class and, hence, at most ν members.

For the problem at hand, $A=\{1, 2, \cdots, 2n\}$ and $T(A)=\{1, 3, 5, \cdots, 2n-1\}$, so that $\nu=n$.

Also solved by J. C. Abad, R. C. Albert, E. B. Anders, Joseph Arkin, M. A. Bershad, Dennis Bertholf, M. G. Beumer (Netherlands), Walter Bluger, Carolyn Brauer, J. L. Brown, Jr., E. O. Buchman, J. P. Burling, J. A. Burslem, David Carlson, P. Carragher, Mannis Charosh, John Christopher, Allan Chuck and P. S. Goldstein (jointly), J. P. Clement, Jr., D. I. A. Cohen, Robert Cohen, J. C. Deel, J. F. Dillon, Jack Dix, D. Z. Djokovic (Yugoslavia), G. E. Engebretsen, Philip Fung, P. K. Garlick, A. A. Gioia, Jerry Goodman, M. Greening (Australia), Alice Guckin, D. M. Hancasky, Edward Hook, R. F. Jackson, R. A. Jacobson, Geoffrey Kandall, B. G. Klein, Kenneth Kramer and Joel Spruck (jointly), S. Lajos (Hungary), E. S. Langford, Kenneth Lebensold, C. C. Lindner, E. L. Magnuson, Andrzej Makowski (Poland), D. C. B. Marsh, Yehoshua Mayer ben-David, Brockway McMillan, M. G. Murdeshwar, Alfrede Neuman, James Nussbaum, R. J. Oberg, Robert Patenaude, C. B. A. Peck, J. M. Perry, Harsh Pittie, J. R. Porter, Donald Quiring, Simeon Reich (Israel), B. E. Rhoades, Daniel Ritchie (Scotland), D. P. Roselle, D. L. Stenson and T. J. Slaukovsky and L. J. Schneider (jointly), C. D. Scudder III, Robert Sibson (England), Robin Sibson (England), D. L. Silverman, R. L. Syverson, M. R. Tiller, A. M. Vaidya, C. Van de Vyle (Belgium), Chinami Wateri (Japan), Lenard Weinstein, R. E. Whitney, C. R. Williams, C. J. York, Theodore Zerger, and the proposer.

Several solvers called attention to Erdös' Problem 3739 [this MONTHLY, 42 (1935) 396, and 44 (1937), 120, (where solutions by Mannis Charosh and by M. Wachsberger and E. Weissfield were published)]. Other references: Paul Erdös, *Note on sequences of integers no one of which is divisible by any other*, Jour. London Math. Soc., 10 (1935) 126–128; R. E. Johnson, *First Course in Abstract Algebra*, Prentice-Hall, (1958) p. 35; Shklarsky, Chentzov, and Yaglom, *U.S.S.R. Olympiad Problem Book*. W. H. Freeman, (1962) p. 169; I. S. Sominskii, *Method of Mathematical Induction*, Blaisdell, (1961) p. 22, problem 27.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before October 31, 1966.

5380. *Proposed by L. P. Comden and S. S. Mitra, Clarkson College of Technology, Potsdam, N. Y.*

Let $f(z)=u(x, y)+iv(x, y)$ be an entire function which is real for real values of z and strictly complex elsewhere, i.e. $v(x, y)\neq 0$ if $y\neq 0$. Prove that $f(z)=az+b$, where $a(\neq 0)$ and b are real. Cf. E 1696 [1965, 550].

5381. *Proposed by V. H. Keiser, University of Colorado*

Let G be an imprimitive group in which all normal subgroups are transitive. Prove that G is nonsolvable.

5382. *Proposed by A. A. Mullin, University of California at Livermore*

Define an arithmetic function μ^* as follows: $\mu^*(1)=1$; $\mu^*(n)=0$ if the mosaic of primes obtained from n has any prime repeated; $\mu^*(n)=(-1)^m$ if the mosaic of n has no prime repeated and m is the number of primes in the mosaic of n. (The mosaic of n is obtained by recursion of the prime-power unique factorization of n upon its own exponents; see problem 5248, [1965, 1140].)

Prove (i): $|\zeta(s)/\zeta(2s)|\leqq\sum_{n=1}^{\infty}|\mu^*(n)/n^s|$, where ζ is Riemann's zeta-function. Prove or disprove (ii) that $\sum_{m\leqq n}\mu^*(m)=o(n)$ entails the Prime Number Theorem.

5383. *Proposed by W. W. Leonard, Susquehanna University, Selinsgrove, Pa.*

A submodule E of a left A-module F is small in F if $E+H=F$ for any submodule H of F implies $H=F$. Prove that Z, the set of integers, is small in Q, the set of rational numbers. Is $Z^{(I)}$ small in $Q^{(I)}$, where I is an infinite set and $Z^{(I)}$ (resp. $Q^{(I)}$) is the direct sum of copies of Z (resp. Q)? All modules are considered as Z-modules.

5384. *Proposed by D. E. Daykin, The University, Kuala Lumpur, Malaya*

Let k be the independence number (coefficient of internal stability) of a graph G of finite order n with no circuits of uneven length. Prove that $\frac{1}{2}\leqq k/n\leqq 1$ and that these bounds are best possible. (See C. Berge, *The Theory of Graphs*, p. 48.)

5385. *Proposed by Stanley Korn, Rensselaer Polytechnic Institute, Troy, N. Y.*

A set S consists of the elements $s_1, \cdots, s_m$, which are related in pairs: $s_i{}^*s_j$, $i\neq j$ (one element may be so related to more than one element, and $s_i{}^*s_j$ is equivalent to $s_j{}^*s_i$). S is to be divided into n nonoverlapping subsets, $n\geqq 1$, (which together contain all the elements of S) such that no two elements in the same subset are *-related. Show that this can always be done if every possible

subset of S consisting of e elements, $e \geqq n$, contains not more than $(n-1)(e-\frac{1}{2}n)$ related pairs.

5386. *Proposed by S. W. Williams, Lehigh University*

Let L be a loop of prime order, S be the set $\{x \mid (ax)b = a(xb);\ a,\ b \in L\}$, and R be the set $\{x \mid (ab)x = a(bx);\ a,\ b \in L\}$. Show that $S = R$.

5387. *Proposed by Michael Gemignani, University of Notre Dame and Saint Mary's College*

Let G be a commutative group with identity 1. If g_1 and g_2 are in G, then define $g_1 \sim g_2$ if $g_1 g_2 = g^2$ for some g in G. $\sim$ is an equivalence relation on G. If G is a finite group of n elements, determine the number of distinct equivalence classes in $G/\sim$.

5388. *Proposed by L. Carlitz, Duke University*

Let p be an odd prime and put

$$K(a, b) = \sum_{x=1}^{p-1} e(ax + bx') \qquad (e(x) = e^{2\pi x i/p}), \quad xx' \equiv 1 \pmod{p},$$

$$J(a) = \sum_{x=1}^{p-1} \psi(x)\psi(x^2 - a),$$

where $\psi(x) = (x/p)$, the Legendre symbol. Show that, for $p \equiv 1 \pmod 4$,

$$\sum_{a=1}^{p-1} K(a, b)J(4a) = \begin{cases} 0 & (\psi(b) = -1) \\ 2pG(c) & (b = c^2,\ c \neq 0), \end{cases}$$

where $G(c) = \sum_{x=1}^{p-1} \psi(x)e(cx)$.

5389. *Proposed by D. J. Newman, Yeshiva University*

Let $f(z)$ be analytic and have $n+1$ zeros in a convex region C. Prove that $\operatorname{Re} f^{(n)}(z)$ vanishes somewhere in C.

SOLUTIONS OF ADVANCED PROBLEMS

Gambler's Ruin

4003 [1941, 483]. *Proposed by G. W. Petrie, South Dakota State School of Mines*

Three men have respectively l, m, and n coins which they match so that the odd man wins. In case all coins appear alike they repeat the throw. Find the average number of tosses required until one man is forced out of the game.

Editorial Note: A solution of this problem by F. Göbel was in the galley proofs of the February issue, but was not printed because of a Mathematical Note by R. C. Read in the same issue. (See 73 (1966) 177–179, and note at foot of p. 179.)

Limit of the nth Power of Matrix

5280 [1965, 428]. *Proposed by T. J. Guglielmo, Chaminade College, Hawaii*

Does the following limit exist and, if so, what is it?

$$\prod_{n=1}^{\infty}\begin{bmatrix}\dfrac{1}{2n} & \dfrac{1}{3n}\\ \dfrac{1}{4n} & \dfrac{1}{5n}\end{bmatrix}.$$

I. *Solution by K. L. Yocom, South Dakota State University.* More generally, if A is any square matrix, then $\prod_{n=1}^{\infty} A/n = \lim_{n\to\infty} A^n/n! = 0$, since the series $\sum_{n=0}^{\infty} A^n/n!$ converges for any A.

II. *Solution by C. G. Cullen, University of Pittsburgh.* The limit is zero and as a matter of fact more is true. The characteristic values of the matrix

$$A = \begin{pmatrix}1/2 & 1/3\\ 1/4 & 1/5\end{pmatrix}$$

are $.35 \pm .33$, less than one in absolute value. Thus $A^n \to 0$ as $n \to \infty$.

Also solved by Claude Andersen, G. Baron & W. Imrich (Austria), D. L. Bruyr, E. O. Buch- man, Orin Chein, L. J. Dickson, D. Ž. Djoković (Yugoslavia), L. S. Evans, S. H. Greene, Harry Guess, Richard Guilfoyle, Eldon Hansen, D. A. Hejhal, Stephen Hoffman, P. G. Kirmser, A.E. Livingston, Charles McCracken, Justin MacCarthy, M. F. Neuts, P. J. Nikolai, H. A. D. Paris (Netherlands), J. C. Parnami (India), J. R. Purdy, R. F. Rinehart, V. L. N. Sarma (India), D. B. Shapiro, Robin Sibson (England), Sidney Spital, Necdet Üçoluk, J. H. van Lint (Netherlands), W. C. Waterhouse, and D. A. Zave.

Identities for Zeros of Hermite Polynomials

5281 [1965, 428]. *Proposed by D. J. Newman, Yeshiva University*

Prove that the general solution of the $N \times N$ algebraic system

$$x_n = \sum_{\substack{i=1\\ i\neq n}}^{N} \frac{1}{x_n - x_i}, \qquad n = 1, 2, \cdots, N$$

is given by the N zeros of $(d/dx)^N e^{-x^2}$ in some order.

Solution by A. S. Householder, Oak Ridge National Laboratory. For a fixed N, let $f(x) = \prod_{i=1}^{N} (x - x_i)$, and let $\phi_n(x) = f(x)/(x - x_n)$. The given equations are

$$(1) \qquad x_n = \phi_n'(x_n)/\phi_n(x_n), \qquad n = 1, 2, \cdots, N.$$

From $\phi_n(x_n) = f'(x_n)$, $2\phi_n'(x_n) = f''(x_n)$ we see that the x_n satisfy $2xf'(x) - f''(x) = 0$.

The left member is a polynomial of degree N and must be a multiple of $f(x)$; thus $f(x)$ satisfies

$$(2) \qquad f'' - 2xf' + 2Nf = 0.$$

The polynomial solution of this differential equation is unique up to a multiplicative constant and is given by

$$e^{x^2}\left(\frac{d}{dx}\right)^N e^{-x^2}.$$

This is the Hermite polynomial $H_n(x)$ except for the factor $(-1)^N$, whose zeros now satisfy (1).

Also solved by Robert Breusch, L. Carlitz, Yu Chang & Sidney Spital, D. Ž. Djoković (Yugoslavia), N. J. Fine, J. H. Halton, M. S. Klamkin, D. Monk (Scotland), D. R. Musser, Stanton Philipp, S. J. Schmahl, J. H. van Lint (Netherlands), Robert Vermes, David Zeitlin, and the proposer.

Editorial Note. Klamkin considers the nature of the more general system

$$F(x_n) = \sum_{i=1, i\neq n}^{N} (x_n - x_i)^{-1}, \qquad n = 1, 2, \cdots, N,$$

$F(x)$ given. If $P(x)=\prod_{i=1}^{N}(x-x_i)$, then $P''(x)-2F(x)P(x)$ is zero at each x_n. The problem of finding $P(x)$ can be handled by the method of undetermined coefficients when $F(x)$ is a rational function.

The Fixed Points of the Bilinear Transformation

5282 [1965, 428]. *Proposed by G. Di Antonio, Fresno State College, California*

Given $T=(az+b)/(cz+d)$, $ad\neq bc$, and let I: $|cz+d|=1$, I': $|cz-a|=1$, be the isometric circles for T and T^{-1}, respectively. It is known (Ford, *Automorphic Functions*, Ch. I) that T is an inversion in I plus a reflection on L, the radical axis of I and I', for the nonloxodromic case. With this information show that, for the hyperbolic case, I and I' each contain a fixed point, and that each fixed point is a limit point of a nested sequence of successively transformed circles starting with I in one case and with I' in the other.

Solution by Simon Vatriquant, Brussels, Belgium. The isometric circles, having the same radius $1/|c|$, their radical axis L is the perpendicular bisector of the line of their centers. If we transform the circle I by T, the reflection on L gives the circle I' and the inversion in I gives a circle C_1 smaller than I and interior to it, with its center on II'. If we repeat the same operation with C_1 we obtain a circle C_2. A new repetition will give a circle C_3, etc., and we have $C_n=T(C_{n-1})$. The sequence $I, C_1, \cdots, C_n, \cdots$ has a limit point A. For, if xy, $x_1y_1, \cdots$, $x_ny_n, \cdots$ are the diameters of $I, C_1, \cdots, C_n, \cdots$ situated on II' all the points $x, x_1, \cdots, x_n, \cdots$ are monotone in the sense II' and the points y, $y_1, \cdots, y_n, \cdots$ are monotone in the sense $I'I$. All the x are on the same side with respect to the y and, conversely, all the y are on the same side with respect to the x. The distance x_ny_n becomes as small as desired. Consequently, there exists one and only one point A separating the x from the y and $\lim x_n=\lim y_n=\lim C_n'=A$.

A may be considered as a point-circle and $A=T(A)$. Thus A is a double-point of T.

Starting from I', with the inverse transformation T^{-1}, we obtain B, the second double-point of T. The midpoint of AB is the midpoint of II'.

The situation is not restricted to the hyperbolic case.

Also solved by the proposer.

Bounds for arctan x- arctan y

5283 [1965, 428]. *Proposed by D. S. Mitrinović, University of Belgrade, Yugoslavia*

Prove the inequality

$$\frac{b-a}{(a^2+r^2)^{1/2}(b^2+r^2)^{1/2}} < \frac{1}{r}\left(\arctan\frac{b}{r} - \arctan\frac{a}{r}\right),$$

where $a<b$ and $r\neq 0$ (arc tan x designates the principal value).

Solution by M. F. Neuts and M. N. Tata, Purdue University. Without loss of generality we take $r>0$. Let points A and B have coordinates (a, r), (b, r), respectively. Then

$$\theta = \arctan\frac{b}{r} - \arctan\frac{a}{r}$$

is the angle AOB. Twice the area of the triangle AOB is

$$r\left| AB\right| = \left| OA\right| \cdot \left| OB\right| \cdot \sin\theta < \left| OA\right| \cdot \left| OB\right| \cdot\theta$$

which is the stated inequality.

Also solved by R. P. Anderson, John Beidler, Robert Bowen, Robert Breusch, C. A. Bridger, R. G. Buschman, Yu Chang & Sidney Spital, G. Di Antonio, D. Ž. Djoković (Yugoslavia), W. O. Egerland, Harry Guess, Stephen Hoffman, H. H. Hunt, A. W. Johnson, Jr., M. S. Klamkin, E. S. Langford, A. E. Livingston, Charles McCracken, L. D. Meeker, Norman Miller, S. N. Mukherjee & V. L. N. Sarma (India), Stanton Philipp, Joel Pitcairn, J. M. Quoniam (France), S. U. Rangarajan (England), P. A. Scheinok, E. J. Schmahl, Marlow Sholander, W. T. Smythe, Necdet Üçoluk, Simon Vatriquant (Belgium), Chinami Watari (Japan), W. C. Waterhouse, J. M. Wild, Jr., and the proposer.

Editorial Note. Many solvers used the expansion formula for sin $(\alpha-\beta)$ to obtain the result. In this way Rangarajan obtains an inequality in the other direction if the angle included does not exceed $\pi/2$. Thus, if $x>y$,

$$\frac{x-y}{(1+x^2)^{1/2}(1+y^2)^{1/2}} < \arctan x - \arctan y < \frac{\pi}{2}\,\frac{x-y}{(1+x^2)^{1/2}(1+y^2)^{1/2}}$$

with the right inequality requiring the additional condition $xy>-1$.

A Nonlinear; Ordinary Differential Equation in Several Variables

5284 [1965, 429]. *Proposed by D. S. Mitrinović, University of Belgrade, Yugoslavia*

Determine all functions $X_k(k=1, 2, \cdots, n)$ each of which depends on a

Starting from I', with the inverse transformation T^{-1}, we obtain B, the second double-point of T. The midpoint of AB is the midpoint of II'.

The situation is not restricted to the hyperbolic case.

Also solved by the proposer.

Bounds for arctan x- arctan y

5283 [1965, 428]. *Proposed by D. S. Mitrinović, University of Belgrade, Yugoslavia*

Prove the inequality

$$\frac{b-a}{(a^2+r^2)^{1/2}(b^2+r^2)^{1/2}} < \frac{1}{r}\left(\arctan\frac{b}{r} - \arctan\frac{a}{r}\right),$$

where $a<b$ and $r\neq 0$ (arc tan x designates the principal value).

Solution by M. F. Neuts and M. N. Tata, Purdue University. Without loss of generality we take $r>0$. Let points A and B have coordinates (a, r), (b, r), respectively. Then

$$\theta = \arctan\frac{b}{r} - \arctan\frac{a}{r}$$

is the angle AOB. Twice the area of the triangle AOB is

$$r|AB| = |OA|\cdot|OB|\cdot\sin\theta < |OA|\cdot|OB|\cdot\theta$$

which is the stated inequality.

Also solved by R. P. Anderson, John Beidler, Robert Bowen, Robert Breusch, C. A. Bridger, R. G. Buschman, Yu Chang & Sidney Spital, G. Di Antonio, D. Ž. Djoković (Yugoslavia), W. O. Egerland, Harry Guess, Stephen Hoffman, H. H. Hunt, A. W. Johnson, Jr., M. S. Klamkin, E. S. Langford, A. E. Livingston, Charles McCracken, L. D. Meeker, Norman Miller, S. N. Mukherjee & V. L. N. Sarma (India), Stanton Philipp, Joel Pitcairn, J. M. Quoniam (France), S. U. Rangarajan (England), P. A. Scheinok, E. J. Schmahl, Marlow Sholander, W. T. Smythe, Necdet Üçoluk, Simon Vatriquant (Belgium), Chinami Watari (Japan), W. C. Waterhouse, J. M. Wild, Jr., and the proposer.

Editorial Note. Many solvers used the expansion formula for sin $(\alpha-\beta)$ to obtain the result. In this way Rangarajan obtains an inequality in the other direction if the angle included does not exceed $\pi/2$. Thus, if $x>y$,

$$\frac{x-y}{(1+x^2)^{1/2}(1+y^2)^{1/2}} < \arctan x - \arctan y < \frac{\pi}{2}\,\frac{x-y}{(1+x^2)^{1/2}(1+y^2)^{1/2}}$$

with the right inequality requiring the additional condition $xy>-1$.

A Nonlinear; Ordinary Differential Equation in Several Variables

5284 [1965, 429]. *Proposed by D. S. Mitrinović, University of Belgrade, Yugoslavia*

Determine all functions $X_k(k=1, 2, \cdots, n)$ each of which depends on a

single variable x_k in such a manner that

$$\left(\sum_{k=1}^{n} X_k\right)\left(\sum_{k=1}^{n} X_k''\right) = \sum_{k=1}^{n} (X_k')^2,$$

with $X_k' = dX_k/dx_k$, $X_k'' = d^2X_k/dx_k^2$.

Solution by Robert Breusch, Amherst College. If $n>1$ and $r\neq s$, differentiation with respect to x_r yields

(A) $$X_r' \cdot \sum X_k'' + X_r''' \cdot \sum X_k = 2X_r'X_r'', \text{ or}$$

(B) $$\sum X_k'' + (X_r'''/X_r')\cdot \sum X_k = 2X_r'', \text{ and similarly}$$

(B′) $$\sum X_k'' + (X_s'''/X_s')\cdot \sum X_k = 2X_s''.$$

Differentiating (A) with respect to x_s we obtain

(C) $$X_r'\cdot X_s''' + X_r'''\cdot X_s' = 0, \quad \text{or} \quad X_r'''/X_r' + X_s'''/X_s' = 0.$$

Adding (B) and (B′) we obtain

(D) $$\sum X_k'' = X_r'' + X_s''.$$

(i) If $n>2$, (D) shows, via $X_1''+X_2''=X_1''+X_3''=X_2''+X_3''$, that $X_1''=X_2''=X_3''=\cdots=A$ (a constant); thus $nA=2A$ and $A=0$. Therefore $X_k=B_kx_k+C_k$, and the given equation implies that $\sum B_k^2=0$. Thus the only nonconstant solutions are linear with complex coefficients, and in particular, there are no nonconstant real solutions if $n>2$.

(ii) If $n=2$, then direct calculation, starting from (C), shows that the only nonconstant real solutions are

$$\text{(a)}\quad X_1 = A_1e^{kx_1} + B_1e^{-kx_1} + C$$
$$X_2 = A_2\cos(kx_2) + B_2\sin(kx_2) - C$$

requiring $4A_1B_1 = A_2^2+B_2^2$, and

$$\text{(b)}\quad X_1 = Ax_1^2 + B_1x_1 + C_1$$
$$X_2 = Ax_2^2 + B_2x_2 + C_2$$

requiring $4A(C_1+C_2) = B_1^2+B_2^2$.

(iii) If $n=1$, then $X_1=Ae^{Bx_1}$ follows directly from the given equations.

Also solved by M. S. Klamkin, and by Yu Chang & Sidney Spital

Simple Turing Machine with One Internal Configuration

5285 [1964, 429]. *Proposed by A. J. Goldman, National Bureau of Standards*

Which (total) functions on one variable can be computed by a simple Turing Machine with only one internal configuration? (We refer to the definition of simple Turing Machine given by Davis, *Computability and Unsolvability*.)

Solution by Fred Rosenblum, University of Wisconsin. We have alphabet $S_0, S_1, \cdots, S_n$ and internal configuration q. A pair of quadruples of the form $[qS_iS_jq, qS_jS_kq]$ is the same as qS_iS_jq in effect, so that the second quadruple is superfluous. The largest alphabet we can use is therefore S_0, S_1, S_2, S_3. For, suppose we have used S_2, S_3 in our quadruples and we want to introduce S_4. S_2 is introduced by qS_0S_2q or qS_1S_2q (in order that S_2 has a chance of being on the tape, otherwise S_2 is not effectively in the alphabet); and S_3 by qS_0S_3q, qS_1S_3q, or qS_2S_3q. But the last quadruple would be of the superfluous type, for we would have $[qS_0S_2q, qS_2S_3q]$ or $[qS_1S_2q, qS_2S_3q]$. Therefore we must have $\{qS_0S_2q, qS_1S_3q\}$ or $\{qS_0S_3q, qS_1S_2q\}$ so that S_4 can be introduced only by qS_3S_4q which would be a superfluous quadruple. Thus we need only S_0, S_1, S_2, S_3 in the alphabet, so that machines can contain at most three quadruples. (A machine with four quadruples would never stop.)

The initial configuration for a single variable x is $qS_1S_1^x$. If a machine T does not contain a quadruple beginning with qS_1, then T computes $f(x)=x+1$, since we arrive at the value by counting S_1's in the final configuration. Such a machine is $\{qS_0S_1q\}$. All further machines will contain qS_1 quadruples. $\{qS_1S_1q\}$ does not stop; any machine containing qS_iS_iq either does not stop or can be represented with that quadruple, so we eliminate all machines containing qS_iS_iq from consideration. $\{qS_1Rq\}$ and $\{qS_1Lq\}$ compute $f(x)=x+1$. $\{qS_1S_0q\}$ and $\{qS_1S_2q\}$ compute $f(x)=x$. ($\{qS_1S_2q\}$ and $\{qS_1S_3q\}$ are the same machines.)

We now consider two quadruple machines. Any machine $\{qS_0Aq, qS_1Bq\}$ where neither A nor B is S_2, will not stop; any machine $\{qS_0S_2q, qS_1Aq\}$ computes the same function as $\{qS_1Aq\}$; any machine $\{qS_1Aq, qS_2Bq\}$, where A is not S_2, computes the same function as $\{qS_1Aq\}$; any machine $\{qS_0Aq, qS_1S_2q\}$ computes $f(x)=x$; $\{qS_1S_2q, qS_2Rq\}$ computes $f(x)=0$; $\{qS_1S_2q, qS_2S_1q\}$ does not stop; and $\{qS_1S_2q, qS_2Aq\}$ computes $f(x)=x$ for $A=S_0, L$.

Finally we consider three quadruple machines. As noted, such a machine must contain the pair P: $\{qS_0S_2q, qS_1S_3q\}$ or Q: $\{qS_0S_3q, qS_1S_2q\}$. But P and Q are no different in effect. $P\cup\{qS_3S_1q\}$ does not stop; $P\cup\{qS_3Rq\}$ computes $f(x)=0$; and $P\cup\{qS_3Aq\}$ computes $f(x)=x$ for $A=S_0, S_2, L$.

We have exhausted all machines and computed only the functions 0, x $x+1$.

Also solved by Carl P. Pixley.

Representations of Square-free Integers

5286 [1965, 429]. *Proposed by Ralph Greenberg, University of Pennsylvania*

Let k be an integer. Prove the existence of infinitely many primes p for which $p+k$ is square-free. Find an asymptotic formula for the number of such primes under a given limit.

I. *Solution by Wolfgang Schwarz, University of Freiburg, Germany.* We prove the theorem: *Let k be a fixed integer. Then for $x\to\infty$,*

$$S = \sum_{\substack{p \leqq x \\ (p+k)\,\text{squarefree}}} 1 = \prod_{p \nmid k} \left(1 - \frac{1}{p(p-1)}\right) \cdot \mathrm{li}\, x + o(x \cdot (\log x)^{-A}),$$

for every $A > 0$.

The proof follows a method given by T. Estermann in J. London Math. Soc., 6 (1931), 219–221. We have

$$S = \sum_{p \leqq x} \mu^2(x+k) = \sum_{p \leqq x} \sum_{m^2 \mid (p+k)} \mu(m) = \sum_{m^2 \leqq (x+k)} \mu(m) \sum_{\substack{p \leqq x \\ p \equiv -k \pmod{m^2}}} 1 = S_1 + S_2 + S_3;$$

where S_1, S_2, S_3 represents a partitioning of the sum S. By the prime number theorem of Page-Walfisz-Siegel (K. Prachav, *Primzahlverteilung*, p. 144, Theorem 8.2) we have for any $A > 0$,

$$\begin{aligned} S_1 &= \sum_{\substack{m \leqq \log^A x \\ (m,k)=1}} \mu(m) \sum_{\substack{p \leqq x \\ p \equiv -k \pmod{m^2}}} 1 \\ &= \mathrm{li}\, x \cdot \sum_{\substack{m \leqq \log^A x \\ (m,k)=1}} \frac{\mu(m)}{\phi(m^2)} + O(x \cdot \log^A x \cdot \exp(-\gamma\sqrt{\log x})) \\ &= B_k \cdot \mathrm{li}\, x + O(x \cdot (\log x)^{-A/2}), \end{aligned}$$

where

$$B_k = \sum_{\substack{m=1 \\ (m,k)=1}}^{\infty} \frac{\mu(m)}{\phi(m^2)} = \prod_{p \nmid k} \left(1 - \frac{1}{p(p-1)}\right) > 0.$$

Clearly

$$S_2 = \sum_{\substack{m \leqq \sqrt{(x+k)} \\ (m,k) > 1}} = O(\sqrt{x}),$$

$$S_3 = \sum_{\substack{\log^A x < m < \sqrt{(x+k)} \\ (m,k)=1}} \frac{\mu(m)}{\phi(m^2)} = O\left(\sum_{\log^A x < m < \sqrt{(x+k)}} \left(\frac{x}{m^2} + 1\right)\right) = O(x \cdot (\log x)^{-A}).$$

The theorem now follows by combining the estimates for S_1, S_2, S_3.

II. *Solution by Louis Comtet, Viroflay, France.* The required asymptotic formula appears in L. Mirsky, *The number of representations of an integer as the sum of a prime and a k-free integer*, this MONTHLY, 56(1949); see Theorem 2, p. 19.

Also solved by Robert Breusch, J. H. van Lint (Netherlands), and the proposer.

Reducible Covers of a Topological Space

5287 [1965, 429]. *Proposed by Michael Gemignani, University of Notre Dame*

Prove or disprove: Let X be a topological space. Then any covering of X by open sets contains a subcovering C' which is minimal in the sense that if any member of C' is removed, then C' no longer covers X.

Solution by several contributors. The solution of this problem is implied immediately by the result of problem 5062 [1963, 1109]: *A topological space X is bicompact if and only if every cover of X has an irreducible subcover.* Counter-examples to the given conjecture are easily found.

Also solved by Steve Armentrout, Robin Ault, G. Baron and W. Imrich (Austria), John Beidler, O. L. Bierman, W. R. Boland, W. H. Bonney, Robert Bowen, R. D. Bronson, D. L. Bruyr, E. O. Buchman, T. J. Burke, George Cain, Jr., William Caldwell, P. L. Chabot, Orin Chein, S. E. Clauss, W. G. Dotson, Jr., Michael Edelstein, E. W. Ewing, J. H. Foster, Harry Gonshor, L. C. Grove, R. W. Heath, G. A. Heuer & S. M. Robinson, K. G. Johnson, R. F. Jolly, James Joseph, P. M. Kannon, Omar Khayyam, Jr., B. G. Klein, H. E. Lahmann (Germany), W. M. Lambert, Jr., H. C. Lauer, L. D. Lavallee, E. B. Leach, M. J. Lempel, Y. F. Lin, R. L. McKinney, J. C. Matthews, M. D. Mavinkurve (India), J. C. Morgan II, N. S. Natarajan (India), Barbara L. Osofsky, W. H. Patterson, Jr., Stephen Portnoy, B. V. Rao & S. B. Rao (India), Ron Rietz, S. M. Robinson, Fred Rosenblum, Francis Sandomierski, V. L. N. Sarma (India), G. F. Schumm, H. L. Shapiro, J. M. Sherman, Robin Sibson (England), J. L. Sieber, W. R. Smythe, Jr., William Stager, U. B. Tewari (India), Gomer Thomas, J. P. Thomas, C. Watari (Japan), W. C. Waterhouse, R. C. Weger, Robert Whiteley, M. R. Wiscamb, S. V. Witt, D. M. Yasnyi, D. A. Zave, Philip Zipse, and the proposer.

Partial Sums of $(C-1)$-summable Series

5288 [1965, 429]. *Proposed by R. F. Pavley and S. I. Mack, Radio Corporation of America, Moorestown, N. J.*

Does there exist a continuous real-valued function f and a real number x such that no subsequence of partial sums of the Fourier series of f, for this x, converges to $f(x)$?

Solution by Chinami Watari, Tôhoku University, Japan. A negative answer to the proposed question is an immediate corollary of the following proposition: *If a series $\sum a_n$ with terms tending to* 0 *is summable to* 0 *by the method of the first arithmetic means, then there is a subsequence $\{s_{n_k}\}$ of the sequence of the partial sums $\{s_n\}$ convergent to* 0.

In fact, if the sequence $\{s_n\}$ contains no subsequence convergent to 0, then there is a positive constant d and a natural number N such that $|s_n| > d$ for $n > N$. Since a_n tends to 0, there is such a natural number N' that $n > N'$ implies $|a_n| < d$. Write $M = \max(N, N')$. If $s_n > d$ for all $n > M$, s_n cannot be summed to 0; similarly if $s_n < -d$ for all $n > M$. But if $s_m > d$ and $s_n < -d$ for some $m, n > M$, there must be an intermediate partial sum s_k $(M < n < k < m$ for example) of absolute value not exceeding d, since each s_j differs, from one to the next, by a term less than d in absolute value. This contradiction implies the validity of the proposition.

It is a consequence of the well-known theorems of Riemann-Lebesgue and Féjer that the Fourier series of a continuous function satisfies (after subtracting a constant if necessary) the hypothesis of the above proposition.

Also solved by W. H. Bonney, Sidney Spital, and the proposer.

Editorial Note. (1) We observe that the proposition used to resolve the question is true for more general summability methods with the same proof as given above.

(2) Spital notes a result of Men'shov establishing the existence of continuous functions for which no subsequence of the partial sums of the Fourier development converges everywhere. See N. K. Bary, *A Treatise on Trigonometric Series*, I, p. 354. This is, perhaps, another interpretation of the stated problem.

Convergence Factors for Sums of Reciprocals of Primes

5289 [1965, 429]. *Proposed by S. E. Payne, Florida State University*

Let the odd primes be denoted by $q_1, q_2, \cdots$. Let $a_1=1/\phi(q_1)$; $a_n=\{1/\phi(q_n)\}(1-a_{n-1})$ for n greater than 1 (ϕ is the Euler totient). Then show that $\sum_{n=1}^{\infty} a_n=1$.

I. *Solution by J. H. van Lint, Technological University, Eindhoven, Netherlands*

The result as stated is not true. For, $a_n \leqq \frac{1}{2}$ and hence $a_n \geqq 1/2(q_n-1)$, implying the divergence of $\sum a_n$.

II. *Solution by Robert Bowen, University of California. Berkeley.* Let a_n be defined (contrary to the printed definition) by $a_n=(1-a_1-a_2-\cdots-a_{n-1})/\phi(q_n)$. By induction, it follows that

$$1-\sum_{k=1}^{n} a_k=\prod_{k=1}^{n}\left(1-\frac{1}{q_k-1}\right).$$

The divergence of $\sum_{k=1}^{\infty} 1/q_k$ now implies that $\sum a_k=1$.

Also solved by L. Carlitz, N. J. Fine, M. S. Klamkin, A. E. Livingston, and R. Sivaramakrishnan (India).

RECENT PUBLICATIONS AND PRESENTATIONS

EDITE BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: K. O. MAY, University of California, Berkeley, and E. P. VANCE, Oberlin College

A Course of Higher Mathematics, Vol. II (Advanced Calculus). By V. I. Smirnov. Translated by D. E. Brown. Translation edited by I. N. Sneddon. Addison-Wesley, Reading, Mass., 1964. xiv+630 pp. $12.50.

This volume, the second of five, is a translation from the sixteenth (revised) Russian edition (Fizmatgiz, Moscow, 1958). According to the Introduction, the

first Russian edition of this volume, listing J. D. Tamarkin as co-author, was published in 1926. Tamarkin, however, in his review of the first German edition of Volume IV (MR 6 42) credits the second and subsequent volumes to Smirnov alone and gives the original publication date of Volume II as 1933. This confusion is most likely due to the substantial changes that the original work underwent after Tamarkin's departure from the U.S.S.R.

The scope of the volume under review is indicated by the following chapter headings: I Ordinary Differential Equations (76 pp.); II Linear Differential Equations (87 pp.); III Multiple and Line Integrals (134 pp.); IV Vector Analysis and Field Theory (51 pp.); V Foundations of Differential Geometry (48 pp.); VI Fourier Series (79 pp.); VII The Partial Differential Equations of Mathematical Physics (149 pp.). There is also an index.

The author sets forth the material in these chapters lucidly; and, although he occasionally appeals to intuitive arguments, e.g., his discussion of the cross product of two vectors (pp. 303–306), his developments are on the whole mathematically sound. There are no excercises, but a wide variety of detailed and mature examples are provided which clearly indicate the versatility of the procedures which have been presented. These features make this book not only an ideal reference work for teachers, but also a usable text for independent study, especially for students of engineering and physics.

The translation reads smoothly, except for a few places where the idiomatic structure has been followed or where a word has been mistranslated. Fortunately, these vexations are minor and do not obscure the intended mathematical thought.

RODNEY ANGOTTI, SUNY at Buffalo

Periodic Differential Equations. By F. M. Arscott. Macmillan, New York, 1964. x+281 pp. $9.50.

This book is a good introduction to linear second order ordinary differential equations with periodic coefficients. The presentation is clear, well motivated and can be understood with little advanced mathematics. Proceeding from the method of separation of variables in partial differential equations, the author in Chapter 1 motivates periodic differential equations and shows the importance of determining conditions for the existence of special types of periodic solutions of these equations. The next four chapters are devoted to a detailed discussion of the Mathieu functions (periodic solutions of period π and 2π of Mathieu's equation) including qualitative properties, approximation procedures, integral equations and asymptotic properties. Chapter 6 deals with the characteristic exponents and the zones of stability of Mathieu's equation. The remaining four chapters treat the same questions in less detail for Hill's equation, the spheroidal wave equation, Lamé's equation and the ellipsoidal wave equation.

J. K. HALE, Brown University

The Elements of Real Analysis. By Robert G. Bartle. Wiley, New York, 1964. 447 pp. $10.95.

This text provides a good introduction to real analysis at the advanced calculus level. The following extracts from the table of contents give a representative picture of the type of material covered. 1. The Real Numbers: fields, supremum principle, Cantor set. 2. The Topology of Cartesian Spaces: Bolzano-Weierstrass theorem, connectedness, compactness, Lebesgue covering theorem, Baire's theorem. 3. Convergence: Cauchy criterion, norm of a function, limit superior, Cesàro summation. 4. Continuous functions: fixed point theorem for contractions, theorems of Stone-Weierstrass, Tietze, and Arzelà-Ascoli, semicontinuity. 5. Differentiation: directional derivative, open mappings, implicit function theorem, extrema with constraints. 6. Integration: Riemann-Stieltjes integral, Riesz representation theorem, transformation of multiple integrals, improper integrals, dominated convergence theorem. 7. Infinite series: Cauchy products, rearrangements, uniform convergence, Tauber's theorem. There is no discussion of Lebesgue integration, line or surface integrals, complex analysis, or Fourier series.

The book contains over 500 exercises, most of them stressing proofs and counter examples. Among these are 33 connected sets of exercises called "projects" which either extend certain ideas discussed in the text or develop specific topics not treated in the text. For example, there are projects on metric spaces, the development of the logarithm as an integral, the Gamma function and the Laplace transform, infinite products, complex power series. Hints and partial solutions are given for selected exercises. The index is unusually complete and contains a number of private jokes.

In summary, the book is well-planned, the level of rigor is high, the notation is well-chosen, and the exposition is clear. It is a welcome addition to the literature.

T. M. APOSTOL, California Institute of Technology

Introduction to Vector and Tensor Analysis. By Robert C. Wrede. Wiley, New York, 1963. 418 pp. $9.75.

This book is the outgrowth of notes developed by the author while teaching various courses at San Jose State College. Its five chapters deal with vector algebra, differentiation of vector-valued functions, partial differentiation, integration (including Green's theorems) and tensor analysis. Nearly all of the treatment is 3-dimensional, with the exception of the final chapter. There are two sections devoted to special and general relativity theory, and some other applications are given in examples.

The book has limited success in its attempt to be logical. A function is defined (Definition 2-1.1) in the modern fashion as a collection of ordered pairs. But even the examples illustrating the definition do not display ordered pairs. The concept of function is not exploited in defining n-tuples, matrices, vectors (sets of n-tuples which transform properly), or tensors. Theorem 3.6.4 is

false, because continuity of the second order partial derivatives is not assumed in testing for a relative maximum or minimum point of a function of two variables. The function whose value is

$$\frac{x^2y^2}{x^2+y^2} - \frac{1}{8}(x^2+y^2)$$

if $(x, y) \neq (0, 0)$, and 0, if $(x, y) = (0, 0)$ should have a local maximum point at the origin according to the theorem, but it doesn't. Also, the author appears to assume in the proof that a sufficient condition for a local maximum or minimum is that the restriction of the function to a straight line through the point have a local maximum or minimum there.

The author has focussed quite a bit of attention on relativity. There is also a discussion of Maxwell's equations, and a derivation of Kepler's laws. I suspect that engineers, more than mathematicians, will miss higher-dimensional problems. Modern engineering problems frequently involve many more than three variables.

An excellent feature of the book is the series of notes describing the historical development of the theory. The author's style of exposition is clear and easy to follow.

E. B. LEACH, Case Institute of Technology

Elementary Logic. By B. Mates. Oxford University Press, New York, 1965. x+227 pp. $6.00.

Here is an introduction to modern symbolic logic with several unique, and in the reviewer's opinion, highly desirable features. If only every University with an introductory (sophomore or junior) course in modern logic would use a text as excellent as this one! The subject matter is sentential and first-order logic, without and with equality. Thus the book prepares the reader for later work in the mainstream of logical theory, rather than, as is frequently the case in books at this level, an idiosyncratic side stream. A first striking feature of the book is that the pre-logical, philosophical bases of logic are not just skimmed over, but are given careful and exact consideration; we refer to topics like: what logic is, use and mention, variables, sets, object-language and meta-language. Another feature is the clarity and precision found throughout the book which, surprisingly, is usually not much in evidence in a beginning book on logic. A precise definition of interpretation and truth is given shortly after introducing the basic language Z of first-order ("elementary") logic. Translation of natural language into Z is carefully discussed, and pitfalls are noted. A simple natural deduction system for Z is given. Consistency and completeness are proved. Logic with equality is introduced, using Tarski's elegant axiom system (including the axiom found to be redundant by Kalish and Montague). There is a chapter on formalized theories (and definitions), and an excellent outline of the history of logic.

Another unusual feature is that the author does not shy away from stating, and proving, important results such as the completeness theorem. Many of these important results are, indeed, left without complete proof, but this is surely better than leaving the reader with only a provincial and fragmentary view of logic. Among the very rare technical errors in the text we mention only the somewhat misleading remarks on categoricity (pp. 177–178). Within the framework of elementary logic, no axiom system with an infinite interpretation is categorical. No misprints were found!

DONALD MONK, University of Colorado

Selecta Heinz Hopf. Herausgegeben zu seinem 70. Geburtstag von der Eidgenössischen Technischen Hochschule Zürich. Springer-Verlag, Berlin—Heidelberg—New York. 1964. viii+310 pp. DM 39. $9.75.

In this volume, dedicated to Hopf in honor of his seventieth birthday by his colleagues in Zürich, 18 of his most famous papers are reprinted. Their original dates of publication range from 1925 to 1955, and they are chosen from his work in the fields of differential geometry, algebraic topology, algebra, Lie Groups, discontinuous transformation groups, and complex analytic manifolds. There is also a bibliography of Hopf's works, listing one book (with P. Alexandroff) and 69 papers (the last one still in preparation). The editors remark in the foreword that the task of selecting which papers were to appear in this volume was difficult; several very fundamental and original papers had to be omitted.

Hopf is noted among mathematicians as one of the clearest and most beautiful of expositors. Many generations of graduate students working in such fields as algebraic topology, differential geometry in the large, etc., must have gotten tremendous inspiration from reading his papers. Due to the rapid rate at which mathematics is advancing nowadays, the results of most of the papers in this book are now looked on as special cases of various larger and more general theories that are described in the latest advanced textbooks and lecture notes. Thus it is unlikely that future generations of graduate students will turn to these papers except for historical reasons. Their publication at this time does serve a useful purpose, however, in that it reminds us once more of the great originality, profundity, versatility, and expository skill of one of the few truly great mathematicians of our era.

W. S. MASSEY, Yale University

Christian Ludwig Gerling an Carl Friedrich Gauss. Sechzig bisher unveröffentlichte Briefe. By Theo. Gerardy. (Band 5. Arbeiten aus der Niedersächsischen Staats- und Universitätsbibliothek Göttingen.) Vanderhoeck & Ruprecht, Göttingen, 1964. 123 pages. DM 14.

In 1927 an extensive correspondence between Gauss and Gerling was published (a total of 388 letters; 163 by Gauss, 225 by Gerling). Quite recently 60 letters from Gerling to Gauss were discovered and are now offered in print.

A few of the letters deal with geodetic triangulations; fifty-six of them deal primarily with the serious domestic problems which arose between Gauss and his sons Wilhelm and Eugen. Eugen emigrated to the United States (Missouri). Judging by these letters Gauss must have leaned to an extraordinary degree on Gerling for advice and encouragement in the family disputes. Translating passages from the "Einleitung" we read: "In the scientific field the yield is poor; but for the Gauss image the letters are most instructive and of interest as documents of the period. They present Gerling as a devoted and clever adviser of his former teacher (Gerling studied under Gauss in Göttingen). The relation between the correspondents changes gradually from the reverent attitude of student versus teacher to a friendship which lasted to Gauss' death (first letter, Nov. 1813; last letter May 1854; Gauss died 1855). Gerling's advice contributed much toward shaping tolerable relations between Gauss and his sons, and helped prevent deepening of the dissonances. Toward the end of the correspondence the tempest has calmed. After all the unhappiness and anxiety two aging friends exchange news of their children and children's children in North America (Gerling's daughter also emigrated to the United States with her husband). Peace has made its entrance, as Gerling had steadfastly predicted through all crises."

A. J. KEMPNER, Boulder, Colorado

Foundations of the Theory of Algebraic Invariants. By G. R. Gurevich. Translated from the Russian by J. R. M. Radok and A. J. M. Spencer. Noordhoff, Groningen, Holland, 1964. viii+418 pp. $16.25.

The classical theory of algebraic invariants, which was cultivated intensively during the latter part of the nineteenth century, has suffered a severe decline during the present century. The cause of this decline has been attributed to Hilbert's basis theorem and the consequent existence theorems. "But why an encystment of existence theorems should be fatal to any body of mathematics is not clear", as E. T. Bell has remarked. Every student of mathematics should acquaint himself with the theory of invariants. The volume under review provides him with an excellent medium for doing so.

The first chapter presents an exposition of the general concepts of invariant theory with respect to the euclidean, affine and projective groups. It may be read with profit by an undergraduate. The same may be said of the second chapter, which develops tensor algebra from first principles. The extensive utilization of tensor algebra distinguishes Gurevich's book from older books on invariant theory. The following topics are treated in detail: invariant processes, Hilbert's basis theorem, complete systems of invariants, Aronhold's symbolic notation, annihilators, affinors and polyvectors. The proofs are clear and the results are illustrated by many examples. The author has included numerous instructive exercises with answers and hints which enhance the value of the book.

LOUIS WEISNER, University of New Brunswick

Statistics and Experimental Design in Engineering and the Physical Sciences. By Norman L. Johnson and Fred C. Leone. Wiley, New York, 1964. Vol. 1, xvi+523, $10.95, vol. 2, x+399, $11.50.

On the whole, this work, and especially the first volume, is as successful a basis for an applied statistics course for engineers as the reviewer has seen. The books are not intended to compete with more theoretically oriented calculus-level texts. While some discussion of general statistical concepts and some proofs are included, their choice is uneven (e.g., inclusion of proof of the mean square convergence of the Robbins-Monro procedure, but of no implication or discussion of sufficiency except in connection with the Blackwell-Rao technique). At the same time, the brief mention given to some ideas shows greater accuracy than many of the more theoretical texts (e.g., in emphasizing the local character of "bestness" of estimators and the lack of nonasymptotic optimality properties of ML estimators).

The great variety of material is the book's strongest feature. For example, the authors include applications of the beta and Weibull distributions, Kolmogorov-Smirnov techniques, robustness, cost of experimentation considerations, orthogonal polynomials, rotatable designs (with, however, no optimality considerations for these or any other designs except in the simplest sampling context), two-stage procedures, stochastic approximation (with, however, an inadequate description of how to locate the maximum of a regression function). On the other hand, again somewhat at random, in 240 pages in volume 2 on the analysis of variance and related matters, there is no mention of techniques for ranking populations, nor of the effect on significance levels of using the indicated preliminary test of homoscedasticity.

Unfortunately, the beginning of volume 1 is encumbered with notational hurdles. A random variable is denoted by x, and "we wish $F(x)$ *always* to mean the c.d.f. of the random variable x, (whatever x may be) and not just some mathematical function of x." Not exactly convenient compared with, say, F_x, but $F(x)$ has thus been defined as a function of a real variable, usually denoted X; but then, instead of indicating the value of this function by $F(x)(X)$ (compare $F_x(X)$), it is denoted by $F(x)_{x=X}$, "the value taken by the c.d. of x when x is put equal to X." A similar notation and description is used for the density function $p(x)$ (which, "like $F(x)$ represents a *property* of the random variable x, not merely a mathematical function of x. So $p(y)$ and $p(x)_{x=y}$ are not necessarily equal," and indeed y has been used here both as random variable and real number). The confusion is completed by writing $\int p(x)dx$ with x as dummy variable. This picture is not helped by the reversed usage of X and x in almost every other book in probability and statistics.

Nevertheless, the broad choice of text material, problems, tables, and references, lead to the assessment given at the beginning of this review. The one semester allotted by the authors to a course based on volume 1 (with possibly two chapters omitted) seems somewhat ambitious.

J. KIEFER, Cornell University

Packing and Covering. By C. A. Rogers. Cambridge Tracts 54, Cambridge University Press, 1964. 111 pp. $5.50.

By emphasizing density estimates for n-dim nsional spaces (n large) this monograph supplements L. Fejes Toth's 1953 Spring r book on distributions in the plane, on the 2-sphere and in 3 dimensions. By tightening such (predominantly measure theoretical) viewpoints as are sufficient to expound the simplest general results of the subject that have already been discovered, it differs basically from Cassel's encyclopedic 1959 Springer survey, as will be evident from the chapter headings: 1. Packing and covering densities, 2. The existence of reasonably dense packings, 3. The existence of reasonably economical coverings, 4. The existence of reasonably dense lattice packings, 5. The existence of reasonably economical lattice coverings, 6. Packings of simplices cannot be very dense, 7. Packings of spheres cannot be very dense, 8. Coverings with spheres cannot be very economical.

The introduction gives a factual report (dramatic by the scarcity of established precise results) on the known results on lattice packings of spheres (table on p. 3 shows the density of closest lattice packing of a sphere for $2, \cdots, 8$ dimensions and good estimates for $9, \cdots, 12$ dimensions), the lattice packing of convex sets, and the packing and covering of convex bodies (mainly in 2 dimensions).

Chapter 1 is remarkable for introducing the periodic packing density $\delta_p(K)$, and the periodic covering density $\theta_p(K)$, of a bounded measurable subsubset K of $\mathcal{E}_n$ which are obtained by optimizing packing (covering) distributions of translates of K with n independent translational symmetries.

In Chapter 4 the author forms the special n-dimensional lattices $\Lambda(\alpha_1, \cdots, \alpha_{n-1}; \eta)$ with basis $a_1=(\chi, 0, \cdots, 0)$, $a_2=(0, \chi, 0, \cdots, 0), \cdots,$ $a_{n-1}=(0, \cdots, 0, \chi, 0)$, $a_n=(\alpha_1\chi, \alpha_2\chi, \cdots, \alpha_{n-1}\chi, \eta)$ such that $\eta>0$, $\chi>0$, $\chi^{n-1}\eta=1$, $0\leqq\alpha_i<1$ for $1\leqq i<n$, and remarks: It has become clear (to the author) that every lattice with determinant 1 can be approximated, arbitrarily closely, by taking η to be small and choosing $\alpha_1, \alpha_2, \cdots, \alpha_{n-1}$ suitably. This point will be made clear to the reader too in a forthcoming note of A. C. Woods. An averaging process over the lattice $\Lambda(\alpha_1, \cdots, \alpha_{n-1}; \eta)$ serves to prove the existence of reasonably close lattice packings for given convex symmetric bodies.

HANS ZASSENHAUS, The Ohio State University

The Theory of Infinite Series. By P. L. Bhatnagar and C. N. Srinivasiengar. National Publishing House, Delhi, 1964. 194 pp. Rs. 10.00.

This short book is patterned very much on the classical books by Knopp and by Bromwich. The book is self-contained and deals with most of the standard topics covered in a detailed course in infinite series. Chapter headings are the following: Infinite sequences, Infinite series: series of positive terms, Series in general, Derangement of series, Double series, Multiplication of series, Uniform convergence, Real power series, Infinite products, Tannery's theorems, Com-

plex series and products, Transformation . . . into a more rapidly convergent series and Summability processes. While the earlier parts of the book contain exercises which are within the reach of a student, the later ones have few or at times none. The virtue of the book is that it contains the standard material presented clearly with mathematical rigor and as such has strong claims for adoption as a text-book. It appears to be free from mathematical errors, but the following remarks seem justified.

1) On page 172 the authors state "Cesàro's method is easier to deal with than Hölder's method and is capable of being extended to the case when k is not necessarily a positive integer." This seems to imply that Hölder's method can be defined only for positive integral orders, which is not true.

2) The proof of the inclusion of the Cesàro methods in the Abel method (pp. 188–190) is very involved and a direct proof, as is found in Hardy's Divergent Series, would have been better. What is otherwise a carefully written text is marred by a very poor job of proof-correction.

M. S. RAMANUJAN, University of Michigan

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor E. A. Walker, New Mexico State University, has received the 1965 Westhafer Award for Research. The Award carries with it a $500 prize.

Florida State University: Drs. Carl Sikkema, University of Michigan, R. C. Moore, Pennsylvania State University, T. G. Hallam, University of Missouri, W. J. Stiles, Georgia Institute of Technology, and S. J. Lomonaco, Jr., St. Louis University, have been appointed Assistant Professors; Assistant Professors R. W. Gilmer and H. F. Kreimer have been promoted to Associate Professors.

Professor J. J. Urbancek, Illinois Teachers College, Chicago-South, retired with the title of Professor Emeritus in August 1965.

Professor Emeritus E. H. Clark, Hiram College, died on October 24, 1964. He was a charter member of the Association.

Associate Professor Dan Hall, Agricultural & Mechanical College of Texas, died on December 24, 1965. He was a member of the Association for 37 years.

Professor Emeritus L. R. Polan, Alfred University, died on September 15, 1965. He was a member of the Association for 34 years.

Sister Marie Gertrude McNeil, Seton Hill College, died on December 12, 1965. She was a member of the Association for 31 years.

Professor Emeritus Mary E. Wells, Vassar College, died on October 7, 1965. She was a charter member of the Association.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

THE FORTY-NINTH ANNUAL MEETING OF THE ASSOCIATION

The Forty-ninth Annual Meeting of the Mathematical Association of America was held at the Sherman House, Chicago, Illinois, from Wednesday to Friday, January 26–28, 1966, in conjunction with the Annual Meeting of the American Mathematical Society. There were registered 3042 persons, including 1835 members of the Association.

Sessions of the Association were held on Wednesday morning, Thursday morning, and Friday morning and afternoon in the Grand Ballroom of the Sherman House. Presiding officers were Professor Wallace Givens on Wednesday morning, Dr. H. O. Pollak for the Panel Discussion on Problems of Departments of Mathematics on Thursday morning, Professor R. P. Boas on Friday morning and Dean Arthur Grad on Friday afternoon. The Program Committee for the meeting consisted of R. P. Boas, Chairman; Wallace Givens, Arthur Grad, E. H. C. Hildebrandt, F. E. Hohn, and M. E. Shanks.

FIRST SESSION OF THE ASSOCIATION

Symposium on Group Theory in Modern Physics

Group Theory in Atomic and Nuclear Spectroscopy, by Professor B. F. Bayman, University of Minnesota.

If the Hamiltonian of a quantum mechanical system is invariant under a group of transformations, the degenerate eigenstates of the Hamiltonian carry a representation of that group of transformations. Groups leaving atomic or nuclear Hamiltonians invariant are the 3-dimensional rotation-reflection group, and the group of permutations of identical particles. In some systems the Hamiltonian is at least approximately invariant under certain unitary and symplectic groups as well. In each case, the spectroscopist needs a list of the irreducible representations of the group, and its character table. He often needs, in addition, explicit expressions for the matrices that reduce Kronecker products of irreducible representations.

Group Theory in Solid State Physics, by Dr. Melvin Lax, Bell Telephone Laboratories.

The nature of a point group, the nature of a translation group, and the way in which these are combined to form a space group was indicated. The use of groups with 10^{23} elements is avoided by introducing multiplier groups with no more than 48 elements. A linear vibration example was used to show that the Dirac characters of the symmetry group of a dynamical problem can be used to simplify a problem to its dynamical core by exhausting all symmetry information. Subgroup techniques are used to determine the form and number of independent parameters in a matrix or tensor.

Group Theory in Elementary Particle Physics, by Professor Morton Hamermesh, University of Minnesota.

The requirements of relativistic invariance forced physicists to become familiar with the Lorentz group. In recent years, other groups such as SU(3), SU(4), and SU(6), have become important in physics. The main interest of the physicist is in the irreducible representations of the group and the reduction of products of representations. The first provides a guide to the classification of elementary particles and suggests the existence of new undiscovered particles. The second gives the physicist relations between different experiments on collisions of elementary particles.

SECOND SESSION OF THE ASSOCIATION

Panel Discussion on Problems of Departments of Mathematics

Presentations on various assigned topics by Professor R. V. Andree, University of Oklahoma, Professor R. C. Buck, University of Wisconsin, Madison, Professor Joseph

Landin, University of Illinois, Chicago, Professor K. O. May, Carleton College and University of California, Berkeley, Professor Anne F. O'Neill, Wheaton College, Norton, Massachusetts, and Professor G. S. Young, Tulane University.

Professor Andree in first discussing "Computer Science as a Branch of Mathematics" pointed out that the current work in computer science is a far cry from that which was done about fifteen years ago and that modern computer science is very much a branch of mathematics. He suggested that, unless mathematicians become aware of this, we are going to be in about the same position on computer science as we are in statistics and that this will be a great disadvantage to both disciplines. Discussing the second topic, "The Variety of the High School Mathematics Backgrounds of Incoming Freshmen" Professor Andree observed that we seem to have an even larger variety of backgrounds now than we had several years ago since there are so many different programs active in the various high schools, not only in Oklahoma, but in other states as well. He suggested that the new CUPM recommendations may hold some answer to this problem.

Professor Buck discussed the "Increasing Diversity Among the Consumers of Mathematics," and the problem this poses for a department in terms of its divided loyalty. He advocated that a department share in the responsibility for designing and teaching new "service" courses, since otherwise there may not be enough attention given to the long term development of skills and understanding, in favor of more immediate short range goals; e.g. a science department might prefer a course in boolean algebra over a course in probability theory. He also discussed the role of the CUPM publications in providing guidance in selecting among alternatives and, in particular, that of the pale green and dark green reports of the Pregraduate Panel as a source of suggestions for course designers and text writers. Finally, he added the plea that developing departments should not feel their prestige at stake if they do *not* attempt to concentrate wholly upon the mathematics major and the pre-professional population.

Professor Landin based his discussion of "How to Start a New Department" upon his experiences of the past year and a half as head of the Mathematics Department at the University of Illinois in Chicago. The essential activities are: recruitment of faculty, development of curricula, building a library, and acquiring physical facilities. He emphasized the importance of a good location including cultural environment and proximity to other universities. He observed that, of several prerequisites for building a department, a sockful of four-leaf clovers is useful. A part of his talk concerned teacher training, and he discussed that part of his University's program.

Professor May, in discussing "The Undergraduate College," suggested that, in order to exploit the advantages and overcome the disadvantages of the undergraduate college that is not part of a large educational complex, mathematics departments need imaginative and exciting programs. These include the cooperation of groups of colleges, new methods of teaching (abandonment of the old "lecture" method for a combination of a small number of large inspirational lectures, more independent study, tutorials, problem sets, movies, and other self study devices), seminars, experimental courses, undergraduate research projects, faculty study and research projects appropriate to the college (especially scholarly research not requiring membership in the "invisible colleges" of currently fashionable research), and other professional activities. These are the keys to recruiting and keeping appropriate staff.

Professor O'Neill, speaking on "Liberal Arts Colleges for Women," discussed the problems that confront the Department of Mathematics at Wheaton College as being in many respects typical of those that arise at all of the eastern residential colleges for women, with a liberal arts curriculum, and a predominately undergraduate student body. Briefly these are: 1. to keep abreast of changes in pre-college mathematics instruction and update college offerings according to the guidelines of CUPM; 2. to place the entering student in the appropriate mathematics course; 3. to find and keep competent

staff; 4. to challenge the girls and to motivate them to utilize fully their talents in mathematics.

Professor Young, discussing the "Private University," felt that for a department in a growing private university, the major conflict is that between the desire to aid the growth of the university and the desire to build a strong mathematics department. How much influence should computing, engineering, biology, medicine have on the directions of growth of the department? Any? Do the internal needs of mathematics as a field require growth in such directions? Examples were given.

General Discussion by the Panel and the Audience.

Annual Business Meeting of the Association; the Association's Fifth Award for Distinguished Service to Mathematics.

THIRD SESSION OF THE ASSOCIATION

The Mathematical Theory of Optimal Stopping Rules, by Professor H. E. Robbins, Columbia University and University of Minnesota.

On a probability space $(\Omega, \mathfrak{F}, P)$ are given in sequence $\mathfrak{F}_1 \subset \mathfrak{F}_2 \subset \cdots$ of sub-σ-algebras of $\mathfrak{F}$ and a sequence $x_1, x_2, \cdots$ of integrable random variables such that x_n is measurable $(\mathfrak{F}_n)$. A *stopping variable* t is a random variable with values in the set $\{1, 2, \cdots\}$ such that the event $[t=n]$ is in $\mathfrak{F}_n$. Then $x_t = x_{t(\omega)}(\omega)$ is a r.v., and we are interested in the value $V = \sup_t E(x_t)$ taken over the class of all t such that $E(x_t)$ exists. If the sup is attained for some t^* we call t^* *optimal*. For example, if $y_1, y_2, \cdots$ are independent and identically distributed, $\mathfrak{F}_n$ = the σ-algebra generated by $y_1 \cdots$, y_n, and $x_n = \max(y_1, \cdots, y_n) - n$, then if $E|y_1| < \infty$ and $E(y_1^+)^2 < \infty$ it can be shown that the value V is the root α of the equation $E(y_1 - \alpha)^+ = 1$, and that t^* = first $n \geqq 1$ such that $y_n \geqq \alpha$ is optimal. The general theory of such problems involves techniques of martingale theory and the study of the iteration of certain functional operations. There are many applications to the theory of probability and statistical decision, and an endless variety of interesting specific problems which are as yet unsolved.

Function Space Integrals and Nonlinear Equations, by Professor M. D. Donsker, Courant Institute of Mathematical Sciences, New York University.

Who Killed Determinants? by Professor K. O. May, Carleton College and University of California, Berkeley.

The literature on determinants to 1964 was analyzed with respect to quantity, time distribution, purpose, content, and quality. General conclusions were illustrated by a detailed history of a particular topic, the derivatives of a determinant with respect to a parameter.

FOURTH SESSION OF THE ASSOCIATION

Symposium on Recent Advances in Group Theory

Recent Developments in the Theory of Group Characters, by Professor Paul Fong, University of Illinois, Chicago.

The theory of group characters has been of great importance in recent results on finite groups, since characters give a method of studying the global structure of a finite group from local properties of the group. These recent demands put upon characters have brought forth interesting approaches and results in the theory. Of particular interest are the notion of coherence for sets of characters and the theory of blocks of characters. Their value has been convincingly demonstrated by the results already achieved.

Classification Problems, by Professor Daniel Gorenstein, Northeastern University.

The classification problems which were discussed reduce by induction to the following situation: A finite simple group G is given, all of whose proper subgroups have composition factors which

are elements of a specified family F of known simple groups; in general, the problem is to prove that G itself is an element of F (occasionally instead to derive a contradiction). The proof of such theorems divides into two major parts: the aim of the first being to show that a portion of the lattice of proper subgroups of G is identical with that of an element G^* of F; while the second part consists in proving that G^* is the only simple group, up to isomorphism, having a lattice with such a subgroup structure. The first part, to which this talk was devoted, involves a study of the maximal subgroups of G containing a given Sylow p-subgroup for suitable primes p. The techniques depend heavily upon ideas originally developed in the Feit-Thompson paper on groups of odd order. The unsolved problem of classifying, up to composition factors, all finite groups whose orders are not divisible by three was discussed to illustrate in a concrete case what is required.

Recent Developments in Finite Groups, by Professor Michio Suzuki, University of Illinois, Urbana.

The speaker surveyed the recent progress in the theory of finite groups, particularly on the problem of classifying all simple groups. There are several phases in this development. This report emphasized the identification problem. The important class of finite simple groups is the finite analogue of Lie groups, unimodular groups, unitary groups, orthogonal groups, etc. The problem is how to recognize or identify a given group as one of the linear groups. The recent works of many people, including those of W. Feit, G. Higman, D. G. Higman, J. McLaughlin, J. Tits, J. Thompson, and the speaker were discussed.

SPECIAL SESSIONS OF THE ASSOCIATION

Film showings were held in the Grand Ballroom as follows: on Tuesday at 7:15 P.M., "The Analog Computer and its Application to Ordinary Differential Equations" (Kinescope recording, produced in collaboration with Professors R. M. Howe and W. Kaplan by the University of Michigan in B&W), at 7:50 P.M., "Can You Hear the Shape of a Drum? A Lecture by Mark Kac" (a CEM Individual Lectures Production in color), at 9:00 P.M., "Challenge in the Classroom: The Methods of R. L. Moore" (a CEM Individual Lectures Production in color); on Wednesday at 7:15 P.M., "Norbert Wiener—Scientist and Philosopher" (a documentary television program produced by NET in B&W, a kinescope recording), at 7:50 P.M., "The Search for Solid Ground" (a panel discussion on modern logic with Professors M. Kac, J. G. Kemeny, H. Rogers, Jr., and R. M. Smullyan; a Science and Engineering TV Journal-Channel 13 Production in B&W, a kinescope recording), at 9:00 P.M., "Mathematical Induction," two lectures by Leon A. Henkin (produced by the MAA Committee on Production of Films in color); on Thursday, at 7:15 P.M., "Infinite Acres," by Melvin Henriksen (a CEM Animated Calculus Film in color), at 7:23 P.M., "Two Puppet Shows," by Charles and Ray Eames, at 7:35 P.M., "Mathematics of the Honeycomb" (The honeybee as "Mathematician"; a Moody Institute of Science Production in color), at 8:00 P.M., "The Classical Groups As a Source of Algebraic Problems: A Lecture by C. W. Curtis" (a Kinescope recording from the 1965 MAA Summer Meeting; a CEM Individual Lectures Production in B&W), at 9:10 P.M., "Pits, Peaks, and Passes: A Lecture on Critical Point Theory by Marston Morse" (a CEM Individual Lectures Production in color).

MEETING OF THE BOARD OF GOVERNORS

The Board of Governors of the Association met on Tuesday morning and afternoon in the Ruby Room of the Sherman House, with 42 members present.

The Board approved the appointment by President Wilder of the following Nominating Committee for 1966: R. A. Rosenbaum, Chairman; Roy Dubisch and L. J. Paige.

The Board re-elected Professor E. A. Cameron as a member of the Finance Committee of the Association for the four-year term, 1966–1969, and it elected Professor D. E. Christie of Bowdoin College as Governor from the Northeastern Section to fill the

unexpired part of the term ending June 30, 1967, of Professor Howard Eves, who had resigned.

The Board approved an amendment in the regulations governing the Putnam Prize Competition providing that no individual may participate in the competition more than four times.

The Board approved the following schedule of future meetings of the Association: Rutgers—The State University, New Brunswick, New Jersey, August 29–31, 1966; Houston, Texas, January 26–28, 1967; University of Toronto, August 28–30, 1967; San Francisco, California, January 25–27, 1968; University of Wisconsin, Madison, Wisconsin, August 26–28, 1968; New Orleans, Louisiana, January, 1969.

The Board voted to discharge the Joint Committee on the Graduate Program in Mathematics with the thanks of the Association, concurring with similar action taken by the Council of the Society at its meeting on January 23 in Chicago. The Board also voted to endorse the recommendation of the Association's Executive and Finance Committees, asking the President to appoint an ad hoc committee of three to consider what activities, if any, the Association might wish to undertake concerning the graduate program in mathematics.

The Executive Director reported the membership of the Association as 16,934 individual members, an increase of 1155 since the corresponding date last year, 2 corporate members, and 227 academic members.

The Board was informed of the work undertaken by COSRIMS (see page 113 of the January, 1966 issue of the MONTHLY) and voted to express its wholehearted support of the project undertaken by COSRIMS on behalf of the mathematical sciences.

ANNUAL BUSINESS MEETING OF THE ASSOCIATION

The annual Business Meeting of the Association was held on Thursday, January 27, 1966, in the Grand Ballroom of the Sherman House, with President Wilder presiding. The Association's fifth Award for Distinguished Service to Mathematics was made to Professor H. M. Gehman of SUNY at Buffalo. The citation (which appeared on pages 1–2 of the January issue of the MONTHLY) was prepared by Professor A. W. Tucker of Princeton University and read by President Wilder. Professor Gehman, in accepting the Award, expressed his great pleasure at having been voted this Award and thanked the past and present officers and members of the Board of Governors for the valuable support they had given him during his terms as Secretary, Treasurer, and Executive Director of the Association.

The Secretary then announced the results of the balloting for officers in which 2087 votes were cast: Professor E. E. Moise of Harvard University was elected as President-Elect for 1966, Professor G. S. Young of Tulane University as First Vice-President for the two-year term, 1966–1967, and Dr. C. R. Phelps of the National Science Foundation and Professor M. L. Smiley of the University of California, Riverside, as Governors for the three-year term, 1966–1968.

The Secretary then reported on some of the actions taken by the Board of Governors on Tuesday. He expressed the Association's gratitude to Professor W. L. Duren, who retired as Chairman of CUPM on December 14, 1965, for the very effective leadership he had given to CUPM and introduced his successor, Professor Richard D. Anderson. He also expressed special thanks on behalf of the Association to Professor C. B. Allendoerfer who had asked to be relieved as Chairman of CEM for the skill and energy with which he had guided the Association's new and difficult projects in various educational media.

The Secretary conveyed the thanks of the Association to the local members of the Committee on Arrangements for the special efforts required for this, by far the largest gathering of mathematicians held anywhere in this country or overseas. He introduced

Professor Haim Reingold, Chairman of the Committee, who had so successfully coordinated all the activities of the Committee; he also made special mention of the energetic and imaginative efforts with which the publicity for this meeting had been directed by Professor L. R. Wilcox. Finally, he acknowledged with deep appreciation the offer of Sister M. Philip, Chairman of the Department of Mathematics, Rosary College, River Forest, Illinois, to have a number of her students assist with registration, and a similar offer of Mr. David A. Bell and a group of fellow students from Roosevelt University; the help of both these groups proved a great asset in handling the unanticipated large registration.

The Editor of the MONTHLY then announced that he had sufficiently many main articles, mathematical notes, and classroom notes on hand to fill the MONTHLY until the expiration of his term at the end of 1966, so that no such articles could be considered for publication by the present editorial board. The June–July issue of the MONTHLY will contain an announcement as to where papers are to be sent for consideration by the new editorial board.

MEETINGS OF OTHER ORGANIZATIONS

The American Mathematical Society held sessions from Monday, January 24 to Thursday, January 27. The thirty-ninth Josiah Willard Gibbs Lecture was delivered by Professor Martin Schwarzschild of Princeton University on Monday evening at 8:00 P.M., in the Grand Ballroom of the Sherman House, on "Stellar Evolution." The retiring presidential address was delivered by Professor J. L. Doob of the University of Illinois, Urbana, on Wednesday at 3:00 P.M. on "Application to Analysis of a Topological Definition of Smallness of a Set." Professor R. A. Askey of the University of Wisconsin, Madison, delivered an address entitled "Norm Inequalities for Some Orthogonal Expansions," on Wednesday at 2:00 P.M. and Professor George Glauberman of the University of Chicago gave an address on "Sylow 2-Subgroups of Finite Groups" on Thursday at 2:00 P.M.

The 1966 Veblen Prize for Geometry was awarded to Professor Stephen Smale of the University of California, Berkeley, and jointly to Professor Barry Mazur of Harvard University and Professor Morton Brown of the University of Michigan.

ARRANGEMENTS, ENTERTAINMENT AND RECREATION

The Committee on Arrangements for the meeting consisted of Haim Reingold, Chairman; H. L. Alder, Arthur Grad, L. A. Kokoris, Joseph Landin, Josephine J. Mehlberg, Seymour Sherman, G. L. Walker, L. R. Wilcox.

Registration headquarters were located in the George Bernard Shaw Room on Sunday and in the foyer of the Grand Ballroom on the mezzanine floor of the Sherman House on Monday through Friday. Books and other exhibits were maintained in the Exhibit Hall on the mezzanine floor. Accommodations for the meeting were handled by the Chicago Convention Bureau, which arranged for reservations in the Sherman House, the Bismarck, the LaSalle, the Y.M.C.A. and other hotels.

The Illinois Institute of Technology and the University of Illinois, Chicago Circle, were hosts at a tea, to which all members of the Association and the Society were invited, on Wednesday from 4:30 P.M. to 6:00 P.M. on the mezzanine floor of the Sherman House.

HENRY L. ALDER, *Secretary*

ACADEMIC MEMBERS ELECTED INTO THE ASSOCIATION

In accordance with the amendments adopted at the business meeting of the Association at Stillwater on August 30, 1961, the Board of Governors at its meeting in Chicago, Illinois, on January 25, 1966, elected to membership in the Association the ninth set of

applicants for academic members (for election of the other eight sets, see April and November issues for 1962–65). Approval for election to membership was given to the following 8 applicants for academic membership:

University of Bridgeport
University of Cincinnati
Elizabeth Seton College
Fairleigh Dickinson University
University of Missouri at St. Louis
Morehouse College
College of Petroleum and Minerals, Dharan, Saudi Arabia
U. S. Air Force Academy

HENRY L. ALDER, *Secretary*

OFFICERS AND COMMITTEES AS OF FEBRUARY 1, 1966

General Offices: SUNY at Buffalo, Buffalo, New York, 14214
Executive Director: H. M. GEHMAN

OFFICERS

President, R. L. WILDER, University of Michigan (1965–66)
President-Elect, E. E. MOISE, Harvard University (1966)
First Vice-President, G. S. YOUNG, Tulane University (1966–67)
Second Vice-President, A. B. WILLCOX, Amherst College (1965–66)
Editor, F. A. FICKEN, New York University (1962–66)
Secretary, H. L. ALDER, University of California, Davis (1965–69)
Treasurer, H. M. GEHMAN, SUNY at Buffalo (1963–67)
Associate Secretary, RAOUL HAILPERN, SUNY at Buffalo (1963–67)

ADDITIONAL MEMBERS OF THE BOARD OF GOVERNORS

Ex-Presidents

C. B. ALLENDOERFER, University of Washington (1961–66)
R. H. BING, University of Wisconsin, Madison (1965–70)
A. W. TUCKER, Princeton University (1963–68)

Elected Members of the Finance Committee

E. A. CAMERON, University of North Carolina (1966–69)
G. B. PRICE, University of Kansas (1964–67)

Governors at Large

ROY DUBISCH, University of Washington (1964–66)
P. S. JONES, University of Michigan (1965–67)
K. O. MAY, University of California, Berkeley, and Carleton College (1964–66)
C. R. PHELPS, National Science Foundation (1966–68)
R. A. ROSENBAUM, Wesleyan University (1965–67)
M. F. SMILEY, University of California, Riverside (1966–68)

Sectional Governors (July 1, 1963–June 30, 1966)

Allegheny Mountain, P. N. CARPENTER, Grove City College
Indiana, J. E. YARNELLE, Hanover College
Kentucky, L. L. SCOTT, University of Louisville

Metropolitan New York, J. N. EASTHAM, Queensborough Community College
Nebraska, L. K. JACKSON, University of Nebraska
Northern California, IRVING SUSSMAN, University of Santa Clara
Oklahoma, GENE LEVY, University of Oklahoma
Rocky Mountain, W. E. BRIGGS, University of Colorado
Wisconsin, J. V. FINCH, Beloit College

Sectional Governors (July 1, 1964–June 30, 1967)

Kansas, CALVIN FOREMAN, Baker University
Missouri, J. J. ANDREWS, St. Louis University
New Jersey, L. F. MCAULEY, Rutgers—The State University
Northeastern, D. E. CHRISTIE, Bowdoin College
Ohio, WADE ELLIS, Oberlin College
Pacific Northwest, A. T. LONSETH, Oregon State University
Southeastern, J. R. WESSON, Vanderbilt University
Southwestern, HARVEY COHN, University of Arizona
Upper New York State, D. E. KIBBEY, Syracuse University

Sectional Governors, (July 1, 1965–June 30, 1968)

Illinois, J. M. H. OLMSTED, Southern Illinois University
Iowa, D. W. WALL, State University of Iowa
Louisiana-Mississippi, P. K. REES, Louisiana State University
Maryland-D.C.-Virginia, DOROTHY L. BERNSTEIN, Goucher College
Michigan, K. W. FOLLEY, Wayne State University
Minnesota, E. J. CAMP, Macalester College
Philadelphia, EMIL GROSSWALD, University of Pennsylvania
Southern California, R. C. JAMES, Harvey Mudd College
Texas, MARTIN WRIGHT, University of Houston

COMMITTEES OF THE ASSOCIATION

Terms of office of members expire, except where otherwise noted, at the Annual Meeting in January following the last year of service listed below. For temporary committees, no terms of office are listed, since they are automatically discharged at the expiration of the President's term of office, which is the Annual Meeting in January, 1967.

EXECUTIVE COMMITTEE

R. L. WILDER, *Chairman* (1964–67); H. L. ALDER (1965–69), F. A. FICKEN (1962–66), H. M. GEHMAN (1963–67), E. E. MOISE (1966–69), A. B. WILLCOX (1965–66), G. S. YOUNG (1966–67), all *ex officio*.

FINANCE COMMITTEE

R. L. WILDER, *Chairman* (1965–66), *ex officio;* E. A. CAMERON (1966–69), G. B. Price (1964–67), H. L. ALDER (1965–69), *ex officio*, H. M. GEHMAN (1963–67), *ex officio*.

COMMITTEE ON ADVISEMENT AND PERSONNEL

A. B. WILLCOX, *Chairman* (1964–66); K. J. ARNOLD (1966–68), B. H. COLVIN (1964–66), E. A. DAVIS (1964–66), J. S. FRAME (1964–66), C. R. PHELPS (1964–66), W. H. SCHMIDT (1966–68).

COMMITTEE ON AN INTERNSHIP PROGRAM IN MATHEMATICS EDUCATION

ROY DUBISCH, *Chairman;* MAX BEBERMAN, D. E. CHRISTIE, E. E. MOISE.

COMMITTEE ON EARLE RAYMOND HEDRICK LECTURES

HANS RADEMACHER, *Chairman* (1964–66); E. E. FLOYD (1965–67), C. B. MORREY (1966–68).

COMMITTEE ON EDUCATIONAL MEDIA

F. B. WRIGHT, *Chairman* (1965–67), L. W. COHEN (1965–66), E. A. CODDINGTON (1966–68), LEON HENKIN (1965–67), J. H. HLAVATY (1965–67), P. S. JONES (1965–66), J. L. KELLEY (1966–68), C. E. RICKART (1965–67), GEORGE SPRINGER (1966–68), M. H. STONE (1965–66), E. P. VANCE (1964–66).

Advisory Committee for Calculus Films: H. M. MACNEILLE, *Director* (1965–67); J. H. CURTISS (1965–67), S. B. JACKSON (1965–67), C. B. MORREY (1965–66), A. W. TUCKER (1965–66).

Advisory Committee for Individual Lectures: L. B. WILLIAMS, *Director* (1965–67); R. C. BUCK (1965–67), L. W. COHEN (1965–67), BERNARD FRIEDMAN (1965–67) P. D. LAX (1965–67).

Advisory Committee for Level I Films: C. B. ALLENDOERFER, *Director* (1965–67); N. J. FINE (1965–67), LEON HENKIN (1965–67), J. H. HLAVATY (1965–67), P. S. JONES (1965–67), E. P. VANCE (1965–67), G. C. WEBBER (1965–67), F. B. WRIGHT (1965–67).

Advisory Committee for Programed Learning: B. H. GERE, *Director* (1965–67); D. W. BLAKESLEE (1965–67), LEONARD GILLMAN (1965–67), ROBERT KALIN (1965–67), A. B. MEWBORN (1965–67), L. W. SMITH (1965–67).

Subcommittee on Television: P. S. JONES, *Chairman;* C. B. ALLENDOERFER, R. C. FISHER.

COMMITTEE ON HIGH SCHOOL CONTESTS

Terms of office of members of this Committee expire on August 31 of the last year of service listed.

C. T. SALKIND, *Chairman* (1963–66); W. H. FAGERSTROM, *Director*, L. C. DALTON (1965–68), J. M. EARL (1965–68), N. S. MENDELSOHN (1965–68), HANS SAGAN (1964–67), W. H. SCHMIDT (1963–66), E. E. STROCK (1964–67).

COMMITTEE ON INSTITUTES

D. L. THOMSEN, JR., *Chairman* (1964–66); D. K. HARRISON (1966–68), MARK KAC (1964–66), JOSEPH LANDIN (1966–68). E. R. MULLINS (1966–68).

COMMITTEE ON INTERNATIONAL COOPERATION IN MATHEMATICAL EDUCATION

WADE ELLIS, *Chairman;* E. G. BEGLE, W. L. DUREN, JR., MAX KRAMER, G. B. PRICE.

COMMITTEE ON PUBLICATIONS

R. P. BOAS, *Chairman* (1965–66); H. S. M. COXETER (1965–67), C. W. CURTIS (1965–66), MARY P. DOLCIANI (1966–68), WALLACE GIVENS (1965–67), I. I. HIRSCHMAN, JR., (1966–68), IVAN NIVEN (1966–68), R. A. ROSENBAUM (1965–66), OLGA TAUSSKY (1966–68), ROY DUBISCH (1964–68), *ex officio*, F. A. FICKEN (1962–66), *ex officio*, H. M. GEHMAN (1963–67), *ex officio.*

Subcommittee on Carus Monographs: IVAN NIVEN, *Chairman* (1966–68); R. P. BOAS (1965–66), OLGA TAUSSKY (1966–68).

Subcommittee on Ford Awards: R. P. BOAS, *Chairman* (1965–66), *ex officio;* C. W. CURTIS (1966–68), R. P. DILWORTH (1965–67).

Subcommittee on MAA Studies in Mathematics: C. W. CURTIS, *Chairman* (1965–66); H. S. M. COXETER (1965–67), WALLACE GIVENS (1965–67).

Subcommittee on Slaught Papers: I. I. HIRSCHMAN, Jr., *Chairman*, (1966–68); MARY P. DOLCIANI (1966–68), R. A. ROSENBAUM (1965–66), F. A. FICKEN (1966), *ex officio*.

Subcommittee on Miscellaneous Publications: R. P. BOAS, *Chairman* (1966), H. L. ALDER (1966–69), H. M. GEHMAN (1966–67), all *ex officio*.

COMMITTEE ON SECONDARY SCHOOL LECTURERS

H. M. GELDER, *Chairman* (1965–67); D. R. BEY (1965–67), B. H. BISSINGER (1966–68), I. C. GENTRY (1964–66), FRANK HAWTHORNE (1964–66), H. V. HUNEKE (1965–67), MAX KRAMER (1964–66).

COMMITTEE ON SECTIONS

L. E. MEHLENBACHER, *Chairman* (1963–66); E. H. BATHO (Acting for M. E. MUNROE Jan 1–Sept. 15, 1966), E. M. BEESLEY (1965–68), C. M. BRADEN (1963–66), M. E. MUNROE (1964–67), ARNOLD WENDT (1966–69), RAOUL HAILPERN (1963–67), *ex officio*.

COMMITTEE ON THE AWARD FOR DISTINGUISHED SERVICE TO MATHEMATICS

G. S. YOUNG, *Chairman* (1965–66), H. M. GEHMAN (1966–68), G. B. PRICE (1964–67).

COMMITTEE ON THE CHAUVENET PRIZE

E. F. BECKENBACH, *Chairman* (1964–66); SAMUEL EILENBERG (1966–68), V. L. KLEE (1965–67).

COMMITTEE ON THE PUTNAM PRIZE COMPETITION

A. M. GARSIA, *Chairman* (1964–66); J. H. MCKAY, *Director* (1965–67), R. E. GREENWOOD (1965–67), N. D. KAZARINOFF (1966–68).

COMMITTEE ON THE UNDERGRADUATE PROGRAM IN MATHEMATICS

R. D. ANDERSON, *Chairman* (1965–67); R. P. BOAS (1966), C. E. BURGESS (1965–66), L. W. COHEN (1966–68), M. L. CURTIS (1965–67), C. R. DEPRIMA (1965–67), BERNARD FRIEDMAN (1966–68), LEONARD GILLMAN (1965–67), SAMUEL GOLDBERG (1966), A. J. HOFFMAN (1966–68), M. GWENETH HUMPHREYS (1965–66), R. C. JAMES (1965–66), L. J. PAIGE (1965–67), B. J. PETTIS (1966–68), ALEX ROSENBERG (1966–68), R. M. THRALL (1965–67), A. B. WILLCOX (1966–68), G. S. YOUNG (1965–66), E. G. BEGLE, *ex officio*, R. L. WILDER (1965–66), *ex officio*.

Advisory Group on Communications: A. B. WILLCOX, *Chairman* (1966–68); E. G. BEGLE (1966), C. E. BURGESS (1966), C. R. DEPRIMA (1966–67), R. M. THRALL (1966–67).

Panel on College Teacher Preparation: LEONARD GILLMAN, *Chairman* (1966–68); C. E. BURGESS (1966–68), D. W. BUSHAW (1966–68), D. E. CHRISTIE (1966–68), L. W. COHEN (1966–68), MEYER JERISON (1966–68), HERMAN MEYER (1966–68), ALEX ROSENBERG (1966–68), E. H. SPANIER (1966–68), J. H. WELLS (1966–68).

Panel on Mathematics for the Biological, Management and Social Sciences: SAMUEL GOLDBERG, *Chairman* (1963–66); ROBERT BUSH (1963–66), B. P. COHEN (1964–66), LEO KATZ (1964–66), H. L. LUCAS (1966), G. B. PRICE (1966), O. H. SCHMITT (1964–66), MARTIN SHUBIK (1964–66), W. A. SPIVEY (1964–66), T. D. STERLING (1966), G. L. THOMPSON (1963–66), R. M. THRALL (1966).

Panel on Pre-Graduate Training: R. P. BOAS, *Chairman* (1966); LOUIS AUSLANDER (1965–67), P. T. BATEMAN (1965–67), D. W. BUSHAW (1965–67), L. W. COHEN (1959–66), L. A. DWIGHT (1964–66), LEONARD GILLMAN (1965–67), G. E. HAY (1964–66), V. L. KLEE (1965–67), HERMAN MEYER (1965–67), A. B. WILLCOX (1965–67).

Panel on Teacher Training: G. S. YOUNG, *Chairman* (1963–66; E. G. BEGLE (1965–67), MARY FOLSOM (1966–67), C. E. HARDGROVE (1966–68), E. R. KOLCHIN (1966–68), E. E. MOISE (1963–66), GEORGE SPRINGER (1966–68), S. S. WILLOUGHBY (1966–68).

Panel on the Physical Sciences and Engineering: R. J. WALKER, *Chairman* (1966); R. C. BUCK (1966), B. H. COLVIN (1965–66), C. R. DEPRIMA (1966), C. A. DESOER (1965–66), MELBA PHILLIPS (1965–66), H. O. POLLAK (1966), A. H. TAUB (1966), G. M. WING (1966).

COMMITTEE ON VISITING LECTURERS

R. E. GASKELL, *Chairman* (1965–67); E. M. BEESLEY (1966–68), R. D. BOSWELL, JR. (1964–66), W. K. MOORE (1966–68), M. W. POWNALL (1965–67), T. H. SOUTHARD (1965–67), W. L. WILLIAMS (1966–68).

EDITORIAL COMMITTEE ON THE FIFTIETH ANNIVERSARY VOLUME

C. B. ALLENDOERFER, *Chairman*, E. A. CAMERON, K. O. MAY.

JOINT COMMITTEE ON EMPLOYMENT OPPORTUNITIES

Terms of office of members of this committee expire on February 28 of the last year of service listed.

M. L. CURTIS, *Chairman* (1962–66, AMS); A. S. HOUSEHOLDER (1965–66, SIAM), D. R. MORRISON (1964–68, MAA).

JOINT COMMITTEE ON PLACES OF MEETINGS

G. L. WALKER, *Chairman*; H. L. ALDER, H. M. GEHMAN, J. W. GREEN, all *ex officio*.

NOMINATING COMMITTEE FOR 1966

R. A. ROSENBAUM, *Chairman*; ROY DUBISCH, L. J. PAIGE.

EDITORIAL BOARDS OF THE ASSOCIATION

AMERICAN MATHEMATICAL MONTHLY (all terms expire December 31, 1966).
Editor: F. A. FICKEN.
Associate Editors: JOSHUA BARLAZ, J. D. BAUM, A. A. BLANK, J. A. BROWN, HERBERT BUSEMANN, LEONARD CARLITZ, H. S. M. COXETER, J. H. CURTISS, GERTRUDE EHRLICH, HOWARD EVES, RAOUL HAILPERN, MARSHALL HALL, JR., L. M. KELLY, A. E. LIVINGSTON, K. O. MAY, J. R. MAYOR, J. M. H. OLMSTED, R. A. ROSENBAUM, E. P. STARKE, E. P. VANCE, ALBERT WILANSKY, H. S. ZUCKERMAN.

MATHEMATICS MAGAZINE (all terms expire December 31, 1968).
Editor: ROY DUBISCH.
Associate Editors: D. B. DEKKER, LADNOR GEISSINGER (Acting for SAM PERLIS, January 1, 1966 to October 1, 1966), RAOUL HAILPERN, R. E. HORTON, J. H. JORDAN, C. T. LONG, SAM PERLIS, RUTH B. RASMUSEN, H. E. REINHARDT, R. W. RITCHIE, J. M. SACHS, HANS SAGAN, C. T. SANDERS, D. E. THORO, L. M. WEINER.

REPRESENTATIVES OF THE ASSOCIATION

On the AAAS Cooperative Committee on the Teaching of Mathematics and Science:
K. O. MAY (1966–68)
On the American Council on Education:
H. L. ALDER, *ex officio*, R. L. WILDER, *ex officio*.
On the Conference Board of the Mathematical Sciences:
H. L. ALDER, *ex officio*, R. L. WILDER, *ex officio*.
On the Council of the American Association for the Advancement of Science:
RICHARD D. ANDERSON (1965–67), M. M. DAY (1966–68)

On the Governing Council of Mu Alpha Theta:
G. B. PRICE (1964–66)
On the National Research Council:
R. H. BING (July 1, 1965–June 30, 1968)
On the U. S. Commission on Mathematical Instruction:
H. O. POLLAK (July 1, 1962–June 30, 1966), LEONARD GILLMAN (July 1, 1965–June 30, 1969).

REPORT OF THE TREASURER FOR THE YEAR 1965

Following is a summary of the report of Professor H. M. Gehman as Treasurer of the Association for the year 1965. The report has been approved by the Finance Committee and accepted by vote of the Board of Governors. Any member of the Association who wishes a copy of the consolidated report of the Treasurer may obtain one by writing to the Buffalo office of the Association.

The reduction in the General Fund is due to a refund to NSF of amounts charged for indirect costs in past years but not allowed. Some of the publication funds have decreased because of printing charges, but in general the funds of the Association continue to increase.

	January 1, 1965	*December 31, 1965*
ASSETS OF THE ASSOCIATION		
M & T Trust Company, checking account	$244,598	$ 69,231
M & T Trust Company, special account	160,959	113,468
M & T Trust Company, third account	2,405	2,388
Securities at market values	228,352	246,840
	$636,314	$431,928
FUNDS OF THE ASSOCIATION		
Current Fund	5,321	6,928
MATHEMATICS MAGAZINE	601	622
Carus Fund	67,022	71,796
Chace Fund	12,163	5,886
Houck Fund	9,849	9,954
Awards Fund	3,133	1,965
Dunkel Fund	31,292	32,952
Anonymous Fund	4,607	4,655
Greenwood Fund	—	8,350
General Fund	235,852	131,161
	$369,840	$274,273
High School Contests	69	5,636
CEM: Subcommittee on TV	2,909	2,290
1964 Cooperative Summer Seminar	4,225	—
1965 Cooperative Summer Seminar	95,907	4,303
1966 Cooperative Summer Seminar	—	29,568
NSF Grants	160,959	113,468
NSF Contributions	2,405	2,388
	$636,314	$431,928

CALENDAR OF FUTURE MEETINGS

Forty-seventh Summer Meeting, Rutgers, The State University, New Brunswick, New Jersey, August 29–31, 1966.

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN

ILLINOIS, Saint Dominic College, St. Charles, May 13–14, 1966.

INDIANA, Indiana State University, Terre Haute, May 14, 1966.

IOWA

KANSAS

KENTUCKY

LOUISIANA-MISSISSIPPI, Jung Hotel, New Orleans, March 4–5, 1967.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN

MINNESOTA

MISSOURI

NEBRASKA

NEW JERSEY, Ocean Township High School, Oakhurst, May 7, 1966.

NORTHEASTERN

NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.

OHIO

OKLAHOMA-ARKANSAS

PACIFIC NORTHWEST, University of Victoria, Victoria, British Columbia, June 17, 1966.

PHILADELPHIA, Villanova University, Villanova, November 1966.

ROCKY MOUNTAIN, Colorado State University, Fort Collins, May 13–14, 1966.

SOUTHEASTERN, Florida Presbyterian College, St. Petersburg, Spring, 1967.

SOUTHERN CALIFORNIA, San Diego State College, San Diego, March 18, 1967.

SOUTHWESTERN

TEXAS

UPPER NEW YORK STATE, St. Bonaventure University, Olean, May 14, 1966.

WISCONSIN, Wisconsin State University, Eau Claire, May 7, 1966.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, New Brunswick, New Jersey, August 30–September 2, 1966.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Washington State University, Pullman, June 20–24, 1966.

ASSOCIATION FOR COMPUTING MACHINERY, Ambassador Hotel, Los Angeles, August 30–September 1, 1966.

ASSOCIATION FOR SYMBOLIC LOGIC

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

INSTITUTE OF MATHEMATICAL STATISTICS, Rutgers, The State University, New Brunswick, New Jersey, August 30–September 2, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Miami Beach, Florida, June 29, 1966.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Miramar Hotel, Santa Monica, California, May 18–20, 1966.

PI MU EPSILON

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Iowa, Iowa City, May 11–14, 1966. (Symposium on Numerical Analysis.)

THE MATHEMATICS AND STATISTICS SECTION of the NATIONAL REGISTER OF SCIENTIFIC AND TECHNICAL PERSONNEL

Cooperating Societies:

American Mathematical Society
American Statistical Association
Association for Computing Machinery
Association for Symbolic Logic
Biometric Society
Econometric Society
Industrial Mathematics Society
The Institute of Mathematical Statistics
Mathematical Association of America
Operations Research Society of America
Society for Industrial and Applied Mathematics
Society of Actuaries

The American Mathematical Society, at the request of the National Science Foundation, will mail the 1966 reporting forms to the Mathematical and Statistics Section of the National Register. The main objective of the National Register is to provide up-to-date information on the scientific manpower resources of the United States. It is also increasingly valuable to our profession as a source of statistical information.

When you receive a National Register Questionnaire, please fill it out and return it promptly to the Headquarters Offices of the Society at P. O. Box 6248, Providence, Rhode Island 02908.

Harry M. Gehman, Executive Director
Mathematical Association of America

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 4

PART II

Papers in Analysis

Number 11
of the
HERBERT ELLSWORTH SLAUGHT
MEMORIAL PAPERS

APRIL 1966

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. Ficken, Department of Mathematics, New York University, New York, N. Y. 10453.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

THIS IS THE OFFICIAL JOURNAL OF THE

MATHEMATICAL ASSOCIATION OF AMERICA, INC.

(Devoted to the Interests of Collegiate Mathematics)

General Office: SUNY at Buffalo, Buffalo, N. Y. 14214. *Executive Director:* H. M. GEHMAN

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

CONTENTS

CAN ONE HEAR THE SHAPE OF A DRUM?

MARK KAC, The Rockefeller University, New York

To George Eugene Uhlenbeck on the occasion of his sixty-fifth birthday

> "La Physique ne nous donne pas seulement l'occasion de résoudre des problèmes . . . , elle nous fait presentir la solution." H. POINCARÉ.

Before I explain the title and introduce the theme of the lecture I should like to state that my presentation will be more in the nature of a leisurely excursion than of an organized tour. It will not be my purpose to reach a specified destination at a scheduled time. Rather I should like to allow myself on many occasions the luxury of stopping and looking around. So much effort is being spent on streamlining mathematics and in rendering it more efficient, that a solitary transgression against the trend could perhaps be forgiven.

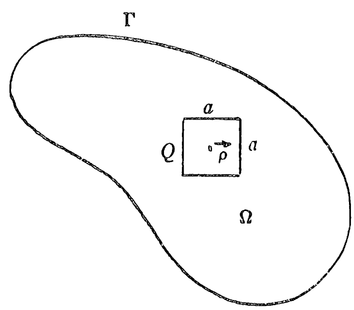


FIG. 1

1. And now to the theme and the title.

It has been known for well over a century that if a membrane Ω, held fixed along its boundary Γ (see Fig. 1), is set in motion its displacement (in the direction perpendicular to its original plane)

$$F(x, y; t) \equiv F(\vec{\rho}; t)$$

obeys the wave equation

$$\frac{\partial^2 F}{\partial t^2} = c^2 \nabla^2 F,$$

where c is a certain constant depending on the physical properties of the membrane and on the tension under which the membrane is held.

I shall choose units to make $c^2 = \frac{1}{2}$.

Of special interest (both to the mathematician and to the musician) are solutions of the form

$$F(\vec{\rho}; t) = U(\vec{\rho})e^{i\omega t},$$

for, being harmonic in time, they represent the *pure tones* the membrane is capable of producing. These special solutions are also known as normal modes.

To find the normal modes we substitute $U(\vec{\rho})e^{i\omega t}$ into the wave equation and see that U must satisfy the equation $\frac{1}{2}\nabla^2 U+\omega^2 U=0$ with the boundary condition $U=0$ on the boundary Γ of Ω, corresponding to the membrane being held fixed along its boundary.

The meaning of "$U=0$ on Γ" should be made clear; for sufficiently smooth boundaries it simply means that $U(\vec{\rho})\to 0$ as $\vec{\rho}$ approaches a point of Γ (from the inside). To show that a membrane is capable of producing a discrete spectrum of pure tones i.e. that there is a discrete sequence of ω's $\omega_1 \leqq \omega_2 \leqq \omega_3 \leqq \cdots$ for which nontrivial solutions of

$$\tfrac{1}{2}\nabla^2 U + \omega^2 U = 0, \qquad U = 0 \text{ on } \Gamma,$$

exist, was one of the great problems of 19th century mathematical physics. Poincaré struggled with it and so did many others.

The solution was finally achieved in the early years of our century by the use of the theory of integral equations.

We now know and I shall ask you to believe me if you do not, that for regions Ω bounded by a smooth curve Γ there is a sequence of numbers $\lambda_1 \leqq \lambda_2 \leqq \cdots$ called eigenvalues such that to each there corresponds a function $\psi(\vec{\rho})$, called an eigenfunction, such that

$$\tfrac{1}{2}\nabla^2 \psi_n + \lambda_n \psi_n = 0$$

and $\psi_n(\vec{\rho})\to 0$ as $\vec{\rho}\to$a point of Γ.

It is customary to normalize the ψ's so that

$$\iint_\Omega \psi_n^2(\vec{\rho})d\vec{\rho} = 1.$$

Note that I use $d\vec{\rho}$ to denote the element of integration (in Cartesian coordinates, e.g., $d\vec{\rho} \equiv dxdy$).

2. The focal point of my exposition is the following problem:

Let Ω_1 and Ω_2 be two plane regions bounded by curves Γ_1 and Γ_2 respectively, and consider the eigenvalue problems:

$\frac{1}{2}\nabla^2 U + \lambda U = 0$ in Ω_1	$\frac{1}{2}\nabla^2 V + \mu V = 0$ in Ω_2
with	with
$U = 0$ on Γ_1	$V = 0$ on Γ_2.

Assume that for each n the eigenvalue λ_n for Ω_1 is equal to the eigenvalue μ_n

for Ω_2. Question: Are the regions Ω_1 and Ω_2 congruent in the sense of Euclidean geometry?

I first heard the problem posed this way some ten years ago from Professor Bochner. Much more recently, when I mentioned it to Professor Bers, he said, almost at once: "You mean, if you had perfect pitch could you find the shape of a drum."

You can now see that the "drum" of my title is more like a tambourine (which really is a membrane) and that stripped of picturesque language the problem is whether we can determine Ω if we know all the eigenvalues of the eigenvalue problem

$$\tfrac{1}{2}\nabla^2 U + \lambda U = 0 \text{ in } \Omega,$$
$$U = 0 \text{ on } \Gamma.$$

3. Before I go any further let me say that as far as I know the problem is still unsolved. Personally, I believe that one cannot "hear" the shape of a tambourine but I may well be wrong and I am not prepared to bet large sums either way.

What I propose to do is to see how much about the shape can be inferred from the knowledge of all the eigenvalues, and to impress upon you the multitude of connections between our problem and various parts of mathematics and physics.

It should perhaps be stated at this point that throughout the paper only *asymptotic properties* of large eigenvalues will be used. This may represent, of course, a serious loss of information and it may perhaps be argued that *precise* knowledge of *all* the eigenvalues may be sufficient to determine the shape of the membrane. It should also be pointed out, however, that quite recently John Milnor constructed two noncongruent sixteen dimensional tori whose Laplace-Betrami operators have exactly the same eigenvalues (see his one page note "Eigenvalues of the Laplace operator on certain manifolds" Proc. Nat. Acad. Sc., 51 (1964) 542).

4. The first pertinent result is that one can "hear" the area of Ω. This is an old result with a fascinating history which I shall now relate briefly.

At the end of October 1910 the great Dutch physicist H. A. Lorentz was invited to Göttingen to deliver the Wolfskehl lectures. Wolfskehl, by the way, endowed a prize for proving, or disproving, Fermat's last theorem and stipulated that in case the prize is not awarded the proceeds from the principal be used to invite eminent scientists to lecture at Göttingen.

Lorentz gave five lectures under the overall title "Alte und neue Fragen der Physik"—Old and new problems of physics—and at the end of the fourth lecture he spoke as follows (in free translation from the original German): "In conclusion there is a mathematical problem which perhaps will arouse the interest of mathematicians who are present. It originates in the radiation theory of Jeans.

"In an enclosure with a perfectly reflecting surface there can form standing

electromagnetic waves analogous to tones of an organ pipe; we shall confine our attention to very high overtones. Jeans asks for the energy in the frequency interval $d\nu$. To this end he calculates the number of overtones which lie between the frequencies ν and $\nu+d\nu$ and multiplies this number by the energy which belongs to the frequency ν, and which according to a theorem of statistical mechanics is the same for all frequencies.

"It is here that there arises the mathematical problem to prove that the number of sufficiently high overtones which lies between ν and $\nu+d\nu$ is independent of the shape of the enclosure and is simply proportional to its volume. For many simple shapes for which calculations can be carried out, this theorem has been verified in a Leiden dissertation. There is no doubt that it holds in general even for multiply connected regions. Similar theorems for other vibrating structures like membranes, air masses, etc. should also hold."

If one expresses this conjecture of Lorentz in terms of our membrane, it emerges in the form:

$$N(\lambda) = \sum_{\lambda_n<\lambda} 1 \sim \frac{|\Omega|}{2\pi}\lambda.$$

Here $N(\lambda)$ is the number of eigenvalues less than λ, $|\Omega|$ the area of Ω and $\sim$ means that

$$\lim_{\lambda\to\infty} \frac{N(\lambda)}{\lambda} = \frac{|\Omega|}{2\pi}.$$

There is an apocryphal report that Hilbert predicted that the theorem would not be proved in his life time. Well, he was wrong by many, many years. For less than two years later Herman Weyl, who was present at the Lorentz' lecture and whose interest was aroused by the problem, proved the theorem in question, i.e. that as $\lambda\to\infty$

$$N(\lambda) \sim \frac{|\Omega|}{2\pi}\lambda.$$

Weyl used in a masterly way the theory of integral equations, which his teacher Hilbert developed only a few years before, and his proof was a crowning achievement of this beautiful theory. Many subsequent developments in the theory of differential and integral equations (especially the work of Courant and his school) can be traced directly to Weyl's memoir on the conjecture of Lorentz.

5. Let me now consider briefly a different physical problem which too is closely related to the problem of the distribution of eigenvalues of the Laplacian.

It can be taken as a basic postulate of classical statistical mechanics that if a system of M particles confined to a volume Ω is in equilibrium with a thermostat of temperature T the probability of finding specified particles at $\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_M$ (within volume elements $\overrightarrow{dr_1}, \overrightarrow{dr_2}, \cdots, \overrightarrow{dr_M}$) is

$$\frac{\exp\left[-\frac{1}{kT}V(\vec{r}_1,\cdots,\vec{r}_M)\right]d\vec{r}_1\cdots d\vec{r}_M}{\int_\Omega\cdots\int_\Omega\exp\left[-\frac{1}{kT}V(\vec{r}_1\cdots\vec{r}_M)\right]d\vec{r}_1\cdots d\vec{r}_M},$$

where $V(\vec{r}_1,\cdots,\vec{r}_M)$ is the interaction potential of the particles and $k=R/N$ with R the "gas constant" and N the Avogadro number.

For identical particles each of mass m obeying the so called Boltzmann statistics the corresponding assumption in quantum statistical mechanics seems much more complicated. One starts with the Schrödinger equation

$$\frac{\hbar^2}{2m}\nabla^2\psi - V(\vec{r}_1,\cdots,\vec{r}_M)\psi = -E\psi \quad \left(\hbar=\frac{h}{2\pi}, \text{ where } h \text{ is the Planck constant}\right)$$

with the boundary condition $\lim\psi(\vec{r}_1.\cdots,\vec{r}_M)=0$, whenever at least one $\vec{r}_k$ approaches the boundary of Ω. (This boundary condition has the effect of confining the particles to Ω.) Let $E_1\leqq E_2\leqq E_3\leqq\cdots$ be the eigenvalues and $\psi_1,\psi_2,\cdots$ the corresponding normalized eigenfunctions. Then the basic postulate is that the probability of finding specified particles at $\vec{r}_1,\vec{r}_2,\cdots,\vec{r}_M$ (within $d\vec{r}_1,\cdots,d\vec{r}_M$) is

$$\frac{\sum_{s=1}^{\infty}e^{-E_s/kT}\psi_s^2(\vec{r}_1,\cdots,\vec{r}_M)d\vec{r}_1\cdots d\vec{r}_M}{\sum_{s=1}^{\infty}e^{-E_s/kT}}.$$

There are actually no known particles obeying the Boltzmann statistics. But don't let this worry you—for our purposes this regrettable fact is immaterial.

Now, let us specialize our discussion to the case of an *ideal* gas which, by *definition*, means that $V(\vec{r}_1,\cdots,\vec{r}_M)\equiv 0$.

Classically, the probability of finding specified particles at $\vec{r}_1,\cdots,\vec{r}_M$ is clearly

$$\frac{d\vec{r}_1\cdots d\vec{r}_M}{|\Omega|^M},$$

where $|\Omega|$ is now the volume of Ω.

Quantum mechanically the answer is not nearly so explicit. The Schrödinger equation for an ideal gas is

$$\frac{\hbar^2}{2m}\nabla^2\psi = -E\psi$$

and the equation is obviously separable.

If I now consider the three-dimensional (rather than the $3M$-dimensional) eigenvalue problem

$$\tfrac{1}{2}\nabla^2\psi(\vec{r}) = -\lambda\psi(\vec{r}), \qquad \vec{r}\in\Omega,$$
$$\psi(\vec{r})\to 0 \quad \text{as} \quad \vec{r}\to \text{the boundary of } \Omega,$$

it is clear that the E_s as well as the $\psi_s(\vec{r}_1, \cdots, \vec{r}_M)$ are easily expressible in terms of the λ's and corresponding $\psi(r)$'s.

The formula for the probability of finding specified particles at $\vec{r}_1, \cdots, \vec{r}_M$ turns out to be

$$\prod_{k=1}^{M} \frac{\sum_{n=1}^{\infty}\exp\left[-\dfrac{\lambda_n\hbar^2}{mkT}\right]\psi_n^2(r_k)}{\sum_{n=1}^{\infty}\exp\left[-\dfrac{\lambda_n\hbar^2}{mkT}\right]}\, d\vec{r}_k.$$

Now, as $\hbar\to 0$ (or as $T\to\infty$) the quantum mechanical result should go over into the classical one and this immediately leads to the conjecture that as

$$\tau\to 0 \qquad \left[\tau = \frac{\hbar^2}{mkT}\right],$$
$$\sum_{n=1}^{\infty} e^{-\lambda_n\tau}\psi_n^2(\vec{r}) \sim \frac{1}{|\Omega|}\sum_{n=1}^{\infty} e^{-\lambda_n\tau}.$$

If instead of a realistic three-dimensional container Ω I consider a two-dimensional one, the result would still be the same

$$\sum_{n=1}^{\infty} e^{-\lambda_n\tau}\psi_n^2(\vec{r}) \sim \frac{1}{|\Omega|}\sum_{n=1}^{\infty} e^{-\lambda_n\tau}, \qquad \tau\to 0,$$

except that now $|\Omega|$ is the area of Ω rather than the volume.

Clearly the result is expected to hold only for $\vec{r}$ in the interior of Ω.

If we believe Weyl's result that (in the two-dimensional case)

$$N(\lambda) \sim \frac{|\Omega|}{2\pi}\lambda, \qquad \lambda\to\infty,$$

it follows immediately by an Abelian theorem that

$$\frac{1}{|\Omega|}\sum_{n=1}^{\infty} e^{-\lambda_n\tau} \sim \frac{1}{2\pi\tau}, \qquad \tau\to 0,$$

and hence that

$$\sum_{n=1}^{\infty} e^{-\lambda_n\tau}\psi_n^2(\vec{r}) \sim \frac{1}{2\pi\tau} = \frac{1}{2\pi}\int_0^{\infty} e^{-\lambda\tau}\, d\lambda.$$

Setting $A(\lambda) = \sum_{\lambda_n<\lambda}\psi_n^2(\vec{r})$, we can record the last result as

$$\int_0^\infty e^{-\lambda\tau}\, dA(\lambda) \sim \frac{1}{2\pi}\int_0^\infty e^{-\lambda\tau}\, d\lambda, \qquad \tau \to 0.$$

Since $A(\lambda)$ is nondecreasing we can apply the Hardy-Littlewood-Karamata Tauberian theorem and conclude what everyone would be tempted to conclude, namely that

$$A(\lambda) = \sum_{\lambda_n<\lambda} \psi_n^2(\vec{r}) \sim \frac{\lambda}{2\pi}, \qquad \lambda \to \infty,$$

for every $\vec{r}$ in the interior of Ω.

Though this asymptotic formula is thus nearly "obvious" on "physical grounds," it was not until 1934 that Carleman succeeded in supplying a rigorous proof.

In concluding this section it may be worthwhile to say a word about the "strategy" of our approach.

We are primarily interested, of course, in asymptotic properties of λ_n for large n. This can be approached by the device of studying the Dirichlet series

$$\sum_{n=1}^\infty e^{-\lambda_n t}$$

for small t. This in turn is most conveniently approached through the series

$$\sum_{n=1}^\infty e^{-\lambda_n t}\psi_n^2(\vec{\rho}) = \int_0^\infty e^{-\lambda t}\, dA(\lambda)$$

and thus we are led to the Abelian-Tauberian interplay described above.

6. It would seem that the physical intuition ought not only provide the mathematician with interesting and challenging conjectures, but also show him the way toward a proof and toward possible generalizations.

The context of the theory of black body radiation or that of quantum statistical mechanics, however, is too far removed from elementary intuition and too full of daring and complex physical extrapolations to be of much use even in seeking the kind of understanding that makes a mathematician comfortable, let alone in pointing toward a rigorous proof.

Fortunately, in a much more elementary context the problem of the distribution of eigenvalues of the Laplacian becomes quite tractable. Proofs emerge as natural extensions of physical intuition and interesting generalizations come within reach.

7. The physical context in question is that of *diffusion theory*, another branch of nineteenth century mathematical physics.

Imagine "stuff," initially concentrated at $\vec{\rho}(\equiv(x_0, y_0))$, diffusing through a plane region Ω bounded by Γ. Imagine furthermore that the stuff gets absorbed ("eaten") at the boundary.

The concentration $P_\Omega(\vec{\rho}\,|\,\vec{r};\,t)$ of matter at $\vec{r}(\equiv(x,\,y))$ at time t obeys the differential equation of diffusion

$$\text{(a)}\qquad \frac{\partial P_\Omega}{\partial t} = \frac{1}{2}\,\nabla^2 P_\Omega,$$

the boundary condition

$$\text{(b)}\qquad P_\Omega(\vec{\rho}\,|\,\vec{r};\,t) \to 0 \text{ as } \vec{r} \text{ approaches a boundary point,}$$

and the initial condition

$$\text{(c)}\qquad P_\Omega(\vec{\rho}\,|\,\vec{r};\,t) \to \delta(\vec{r} - \vec{\rho}) \quad \text{as} \quad t \to 0;$$

here $\delta(\vec{r}-\vec{\rho})$ is the Dirac "delta function," with "value" ∞ if $\vec{r} = \vec{\rho}$ and 0 if $\vec{r} \neq \vec{\rho}$.

The boundary condition (b) expresses the fact that the boundary is absorbing and the initial condition (c) the fact that initially all the "stuff" was concentrated at $\vec{\rho}$.

I have again chosen units so as to make the diffusion constant equal to $\frac{1}{2}$.

As is well known the concentration $P_\Omega(\vec{\rho}\,|\,\vec{r};\,t)$ can be expressed in terms of the eigenvalues λ_n and normalized eigenfunctions $\psi_n(\vec{r})$ of the problem

$$\tfrac{1}{2}\nabla^2\psi + \lambda\psi = 0 \text{ in } \Omega,$$
$$\psi = 0 \text{ on } \Gamma.$$

In fact, $P_\Omega(\vec{\rho}\,|\,\vec{r};\,t) = \sum_{n=1}^{\infty} e^{-\lambda_n t}\psi_n(\vec{\rho})\psi_n(\vec{r})$.

Now, for small t, it appears intuitively clear that particles of the diffusing stuff will not have had enough time to have felt the influence of the boundary Ω. As particles begin to diffuse they may not be aware, so to speak, of the disaster that awaits them when they reach the boundary.

We may thus expect that in some approximate sense

$$P_\Omega(\vec{\rho}\,|\,\vec{r};\,t) \sim P_0(\vec{\rho}\,|\,\vec{r};\,t), \quad \text{as} \quad t \to 0,$$

where $P_0(\vec{\rho}\,|\,\vec{r};\,t)$ still satisfies the same diffusion equation

$$\text{(a}'\text{)}\qquad \frac{\partial P_0}{\partial t} = \tfrac{1}{2}\,\nabla^2 P_0$$

and the same initial condition

$$\text{(c}'\text{)}\qquad P_0(\vec{\rho}\,|\,\vec{r};\,t) = \delta(\vec{r} - \vec{\rho}), \qquad t \to 0,$$

but is otherwise unrestricted.

Actually there is a slight additional restriction without which the solution is not unique (a remarkable fact discovered some years ago by D. V. Widder). The restriction is that $P_0 \geqq 0$ (or more generally that P_0 be bounded from below).

A similar restriction for P_Ω is not needed since for diffusion in a *bounded* region it follows automatically.

An explicit formula for P_0 is, of course, well known. It is

$$P_0(\vec{\rho}\,|\,\vec{r};t) = \frac{1}{2\pi t}\exp\left[-\frac{\|\vec{r}-\vec{\rho}\|^2}{2t}\right],$$

where $\|\vec{r}-\vec{\rho}\|$ denotes the Euclidean distance between $\vec{\rho}$ and $\vec{r}$.

I can now state a little more precisely the principle of "not feeling the boundary" explained a moment ago.

The statement is that as $t \to 0$

$$P_\Omega(\vec{\rho}\,|\,\vec{r};t) = \sum_{n=1}^{\infty} e^{-\lambda_n t}\psi_n(\vec{\rho})\psi_n(\vec{r}) \sim \frac{1}{2\pi t}\exp\left[-\frac{\|\vec{r}-\vec{\rho}\|^2}{2t}\right] = P_0(\vec{\rho}\,|\,\vec{r};t),$$

where $\sim$ stands here for "is approximately equal to." This is a bit vague but let it go at that for the moment.

If we can trust this formula even for $\vec{\rho}=\vec{r}$ we get

$$\sum_{n=1}^{\infty} e^{-\lambda_n t}\psi_n^2(\vec{r}) \sim \frac{1}{2\pi t}$$

and if we display still more optimism we can integrate the above and, making use of the normalization condition

$$\int_\Omega \psi_n^2(\vec{r})\,d\vec{r} = 1,$$

obtain

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t}.$$

We recognize immediately the formulas discussed a while back in connection with the quantum-statistical-mechanical treatment of the ideal gas. If we apply the Hardy-Littlewood-Karamata theorem, alluded to before, we obtain as corollary the theorems of Carleman and Weyl.

To do this, however, we must be allowed to interpret $\sim$ as meaning "asymptotic to."

8. Now, a little mathematical soul-searching. Aren't we as far from a rigorous treatment as we were before? True, diffusion is more familiar than black body radiation or quantum statistics. But familiarity gives comfort, at best, and comfort may still be (and often is) miles away from the rigor demanded by mathematics.

Let us see then what we can do about tightening the loose talk.

First let me dispose of a few minor items which may cause you worry.

When I write $\psi=0$ on Γ or $P(\vec{\rho}\,|\,\vec{r};t)\to 0$ as $\vec{r}$ approaches a boundary point of Ω there is always a question of interpretation.

Let me assume that Γ is sufficiently regular so that no ambiguity arises i.e.

$$P(\vec{\rho}\mid\vec{r};t)\to 0 \quad \text{as} \quad \vec{r}\to \text{a boundary point of } \Omega,$$

means exactly what it says, while $\psi=0$ on Γ means

$$\psi\to 0 \quad \text{as} \quad \vec{r}\to \text{a boundary point of } \Omega.$$

Likewise, $P(\vec{\rho}\mid\vec{r};t)\to\delta(\vec{r}-\vec{\rho})$ as $t\to 0$, has the obvious interpretation, i.e.

$$\lim_{t\to 0}\iint_A P(\vec{\rho}\mid\vec{r};t)\,d\vec{r}=1$$

for every open set A containing $\vec{\rho}$.

Now, to more pertinent items. If the mathematical theory of diffusion corresponds in any way to physical reality we should have the inequality

$$P_\Omega(\vec{\rho}\mid\vec{r};t)\leqq P_0(\vec{\rho}\mid\vec{r};t)=\frac{\exp\left[-\dfrac{\|\vec{\rho}-\vec{r}\|^2}{2t}\right]}{2\pi t}.$$

For surely less stuff will be found at $\vec{r}$ at time t if there is a possibility of matter being destroyed (on the boundary Γ of Ω) than if there were no possibility of such destruction.

Now let Q be a square with center at $\vec{\rho}$ totally contained in Ω. Let its boundary act as an absorbing barrier and denote by $P_Q(\vec{\rho}\mid\vec{r};t)$, $\vec{r}\in Q$, the corresponding concentration at $\vec{r}$ at time t.

In other words, P_Q satisfies the differential equation

$$\text{(a}''\text{)} \qquad \frac{\partial P_Q}{\partial t}=\frac{1}{2}\nabla^2 P_Q$$

and the initial condition

$$\text{(c}''\text{)} \qquad P_Q(\vec{\rho}\mid\vec{r};t)\to\delta(\vec{r}-\vec{\rho}) \quad \text{as} \quad t\to 0.$$

It also satisfies the boundary condition

$$\text{(b}''\text{)} \qquad P_Q(\vec{\rho}\mid\vec{r};t)\to 0 \quad \text{as} \quad \vec{r}\to \text{a boundary point of } Q.$$

Again it appears obvious that

$$P_Q(\vec{\rho}\mid\vec{r};t)\leqq P_\Omega(\vec{\rho}\mid\vec{r};t), \qquad \vec{r}\in Q,$$

for the diffusing stuff which reaches the boundary of Q is lost as far as P_Q is concerned but *need not* be lost as a contribution to P_Ω.

Q has been chosen so simply because $P_Q(\vec{\rho}\mid\vec{r};t)$ is known explicitly, and, in particular

$$P_Q(\vec{\rho}\mid\vec{\rho};t) = \frac{4}{a^2} \sum_{\substack{m,n \\ \text{odd integers}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right],$$

where a is the side of the square.

The combined inequalities

$$P_Q(\vec{\rho}\mid\vec{r};t) \leqq P_\Omega(\vec{\rho}\mid\vec{r};t) \leqq \frac{\exp\left[-\dfrac{\|\vec{r}-\vec{\rho}\|^2}{2t}\right]}{2\pi t}$$

hold for all $\vec{r}\in Q$ and in particular for $\vec{r}=\vec{\rho}$. In this case we get

$$\frac{4}{a^2} \sum_{\substack{m,n \\ \text{odd integers}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right] \leqq \sum_{n=1}^{\infty} e^{-\lambda_n t}\psi_n^2(\vec{\rho}) \leqq \frac{1}{2\pi t}$$

and it is a simple matter to prove that as $t\to 0$ we have *asymptotically*

$$\frac{4}{a^2} \sum_{\substack{m,n \\ \text{odd integers}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right] \sim \frac{1}{2\pi t}\,.$$

Thus asymptotically for $t\to 0$ $\sum_{n=1}^{\infty} e^{-\lambda_n t}\psi_n^2(\vec{\rho})\sim 1/2\pi t$ and Carleman's theorem follows.

It is only a little harder to prove Weyl's theorem.

If one integrates over Q the inequality

$$\frac{4}{a^2} \sum_{\substack{m,n \\ \text{odd}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right] \leqq \sum_{n=1}^{\infty} e^{-\lambda_n t}\psi_n^2(\vec{\rho})$$

one obtains

$$4 \sum_{\substack{m,n \\ \text{odd}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right] \leqq \sum_{n=1}^{\infty} e^{-\lambda_n t} \iint_Q \psi_n^2(\vec{\rho})\, d\vec{\rho}.$$

We now cover Ω with a net of squares of side a, as shown in Fig. 2, and keep only those contained in Ω. Let $N(a)$ be the number of these squares and let $\Omega(a)$ be the union of all these squares. We have

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} = \sum_{n=1}^{\infty} e^{-\lambda_n t} \iint_\Omega \psi^2(\vec{\rho})\, d\vec{\rho} \geqq \sum_{n=1}^{\infty} e^{-\lambda_n t} \iint_{\Omega(a)} \psi_n^2(\vec{\rho})\, d\vec{\rho}$$

$$\geqq 4N(a) \sum_{\substack{m,n \\ \text{odd}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right]$$

and, integrating the inequality $P_\Omega(\vec{\rho}\,|\,\vec{\rho};\ t) \leqq 1/2\pi t$ over Ω we get $\sum_{n=1}^{\infty} e^{-\lambda_n t} \leqq |\Omega|/2\pi t$.

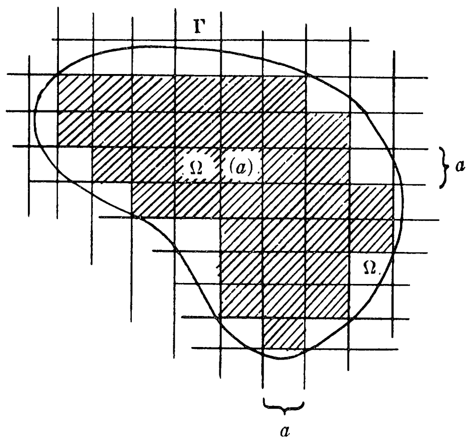


FIG. 2

Noting that $N(a)a^2 = |\Omega(a)|$ we record the fruits of our latest labor in the form of the inequality

$$|\,\Omega(a)\,|\ \frac{4}{a^2} \sum_{\substack{m,n \\ \text{odd}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right] \leqq \sum_{n=1}^{\infty} e^{-\lambda_n t} \leqq \frac{|\,\Omega\,|}{2\pi t}\,.$$

From the fact (already noted above) that

$$\lim_{t\to 0} 2\pi t \frac{4}{a^2} \sum_{\substack{m,n \\ \text{odd}}} \exp\left[-\frac{(m^2+n^2)\pi^2}{2a^2}t\right] = 1$$

we conclude easily that

$$|\,\Omega(a)\,| \leqq \liminf_{t\to 0} 2\pi t \sum_{n=1}^{\infty} e^{-\lambda_n t} \leqq \limsup_{t\to 0} 2\pi t \sum_{n=1}^{\infty} e^{-\lambda_n t} \leqq |\,\Omega\,|\,;$$

and since, by choosing a sufficiently small, we can make $|\Omega(a)|$ arbitrarily close to $|\Omega|$, we must have $\lim_{t\to 0} 2\pi t \sum_{n=1}^{\infty} e^{-\lambda_n t} = |\Omega|$ or, in other words,

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\,\Omega\,|}{2\pi t}, \qquad t \to 0.$$

9. Are we now through with rigor? Not quite. For while the inequalities

$$P_\Omega(\vec{\rho}\,|\,\vec{r};\ t) \leqq \frac{\exp\left[-\dfrac{\|\vec{r}-\vec{\rho}\|^2}{2t}\right]}{2\pi t}$$

$$P_\Omega(\vec{\rho} \mid \vec{r}; t) \geqq P_Q(\vec{\rho} \mid \vec{r}; t), \qquad \vec{r} \in Q,$$

are utterly obvious on intuitive grounds they must be proved. Let me indicate a way of doing it which is probably by far not the simplest. I am choosing it to exhibit yet another physical context.

It has been known since the early days of this century, through the work of Einstein and Smoluchowski, that diffusion is but a macroscopic manifestation of microscopic Brownian motion.

Under suitable physical assumptions $P_\Omega(\vec{\rho} \mid \vec{r};\ t)$ can be interpreted as the probability density of finding a free Brownian particle at $\vec{r}$ at time t if it started on its erratic journey at $t=0$ from $\vec{\rho}$ and if it gets absorbed when it comes to the boundary of Ω.

If a large number N of independent free Brownian particles are started from $\vec{\rho}$ then

$$N \iint_A P(\vec{\rho} \mid \vec{r}; t)\, d\vec{r}$$

is the average number of these particles which are found in A at time t. Since the statistical percentage error is of the order $1/\sqrt{N}$ continuous diffusion theory is an excellent approximation when N is large.

A significant deepening of this point of view was achieved in the early twenties by Norbert Wiener. Instead of viewing the problem as a problem in *statistics* of *particles* he viewed it as a problem in *statistics* of *paths*. Without entering into details let me review briefly what is involved here.

Consider the set of all continuous curves $\vec{r}(\tau)$, $0 \leqq \tau < \infty$, starting from some arbitrarily chosen origin O. Let $\Omega_1, \Omega_2, \cdots, \Omega_n$ be open sets and $t_1 < t_2 < \cdots < t_n$ ordered instants of time. The Einstein-Smoluchowski theory required that (with suitable units)

$$\text{Prob.}\,\{\vec{\rho} + \vec{r}(t_1) \in \Omega_1, \vec{\rho} + \vec{r}(t_2) \in \Omega_2, \cdots, \vec{\rho} + \vec{r}(t_n) \in \Omega_n\}$$
$$= \int_{\Omega_1} \cdots \int_{\Omega_n} P_0(\vec{\rho} \mid \vec{r}_1; t_1) P_0(\vec{r}_1 \mid \vec{r}_2; t_2 - t_1) \cdots P_0(\vec{r}_{n-1} \mid \vec{r}_n; t_n - t_{n-1})\, d\vec{r}_1 \cdots d\vec{r}_n$$

where, as before,

$$P_0(\vec{\rho} \mid \vec{r}; t) = \frac{1}{2\pi t} \exp\left[-\frac{\|\vec{r} - \vec{\rho}\|^2}{2t}\right].$$

Wiener has shown that it is possible to construct a completely additive measure on the space of all continuous curves $\vec{r}(\tau)$ emanating from the origin such that the set of curves $\vec{\rho} + \vec{r}(\tau)$ which at times $t_1 < t_2 < \cdots < t_n$ find themselves in open sets $\Omega_1, \Omega_2, \cdots, \Omega_n$ respectively, has measure given by the Einstein-Smoluchowski formula above.

The set of curves such that $\vec{\rho} + \vec{r}(\tau) \in \Omega$, $0 \leqq \tau \leqq t$, and $\vec{\rho} + \vec{r}(t) \in A$ (A—an open set) turns out to be measurable and it can be shown, if Ω has sufficiently

smooth boundaries, that this measure is equal to

$$\int_A P_\Omega(\vec{\rho} \mid \vec{r}; t)\, d\vec{r}.$$

This is not a trivial statement and it should come as no surprise that it trivially implies the inequalities we needed a while back to make precise the principle of not feeling the boundary.

In fact, as the reader no doubt sees, the inequalities in question are simply a consequence of the fact that if sets $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are such that

$$\mathcal{A} \subset \mathcal{B} \subset \mathcal{C}$$

then meas. $\mathcal{A} \leqq$ meas. $\mathcal{B} \leqq$ meas. $\mathcal{C}$.

One final remark before we go on. The set of curves for which

$$\vec{\rho} + \vec{r}(\tau) \in \Omega, \qquad 0 \leqq \tau \leqq t \quad \text{and} \quad \vec{\rho} + \vec{r}(t) \in A$$

is measurable even if the boundary of Ω is quite wild. The measure can still be written as $\int_A P_\Omega(\vec{\rho} \mid \vec{r}; t) d\vec{r}$ and it can be shown that in the interior of Ω, $P_\Omega(\vec{\rho} \mid \vec{r}; t)$ satisfies the diffusion equation $\partial P_\Omega / \partial t = \frac{1}{2} \nabla^2 P_\Omega$ as well as the initial condition

$$\lim_{t \to 0} \int_A P_\Omega(\vec{\rho}; \vec{r}; t)\, d\vec{r} = 1,$$

for all open sets A such that $\vec{\rho} \in A$.

It is, however, no longer clear how to interpret the boundary condition that

$$P_\Omega(\vec{\rho} \mid \vec{r}; t) \to 0 \quad \text{when} \quad \vec{r} \to \Gamma.$$

This difficulty forces the classical theory of diffusion to consider reasonably smooth boundaries. The probabilistic interpretation of $P_\Omega(\vec{\rho} \mid \vec{r};\, t)$ provides a natural definition of a *generalized solution* of the boundary value problem under consideration.

10. We are now sure that we can hear the area of a drum and it may seem that we spent a lot of effort to achieve so little.

Let me now show you that the approach we used can be extended to yield more, but to avoid certain purely geometrical complications I shall restrict myself to convex drums.

We have achieved our first success by introducing the principle of not feeling the boundary. But if $\vec{\rho}$ is close to the boundary Γ of Ω then the diffusing particles starting from $\vec{\rho}$ will, to some extent, begin to be influenced by Γ.

Let $\vec{q}$ be the point on Γ closest to $\vec{\rho}$ and let $l(\vec{\rho})$ be the straight line perpendicular to the line joining $\vec{\rho}$ and $\vec{q}$. (See Fig. 3.) Then a diffusing particle starting from ρ will see for a short time the boundary Γ as the straight line $l(\vec{\rho})$.

One may say, using again somewhat picturesque language, that, for small t, the particle has not had time to feel the curvature of the boundary.

If this principle is valid I should be allowed to approximate (for small t)

$$P_\Omega(\vec{\rho}\,|\,\vec{r};t) \quad \text{by} \quad P_{l(\vec{\rho})}(\vec{\rho}\,|\,\vec{r};t),$$

where $P_{l(\vec{\rho})}(\vec{\rho}\,|\vec{r};t)$ satisfies again the diffusion equation

$$\frac{\partial P}{\partial t} = \frac{1}{2}\nabla^2 P$$

with the initial condition $P \to \delta(\vec{\rho}-\vec{r})$ as $t \to 0$, but with the boundary condition

$$P_{l(\vec{\rho})}(\vec{\rho}\,|\,\vec{r};t) \to 0 \quad \text{as } \vec{r} \text{ approaches a point on } l(\vec{\rho}).$$

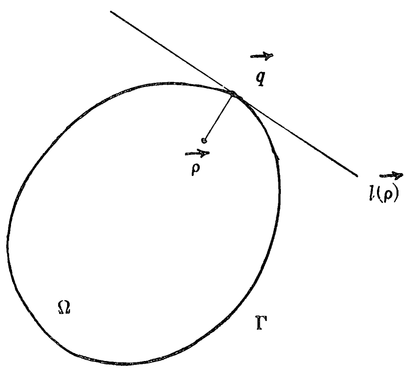


FIG. 3

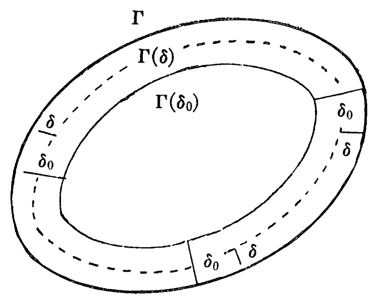


FIG. 4

Carrying this optimism as far as possible we would expect that to a good approximation

$$\int_\Omega P_\Omega(\vec{\rho}\,|\,\vec{\rho};t)\,d\vec{\rho} \sim \int_\Omega P_{l(\vec{\rho})}(\vec{\rho}\,|\,\vec{\rho};t)\,d\vec{\rho}.$$

It is well known that

$$P_{l(\vec{\rho})}(\vec{\rho}\,|\,\vec{\rho};t) = \frac{1 - e^{-2\delta^2/t}}{2\pi t},$$

where $\delta = \|\vec{q}-\vec{\rho}\|$ = minimal distance from $\vec{\rho}$ to Γ. Thus (hopefully!)

$$\iint_\Omega P_\Omega(\vec{\rho}\,|\,\vec{\rho};t)\,d\vec{\rho} = \sum_{n=1} e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t} - \frac{1}{2\pi t}\int_\Omega e^{-2\delta^2/t}\,d\vec{\rho}.$$

Here $|\Omega|/2\pi t$ is our old friend from before and it remains to calculate asymptotically (as $t \to 0$) the integral $\int_\Omega e^{-2\delta/t}d\vec{\rho}$. To do this consider the curve $\Gamma(\delta)$ of points in Ω whose "distance" from Γ is δ. (See Fig. 4.)

For small enough δ, $\Gamma(\delta)$ is well defined (and even convex) and the major contribution to our integral comes from small δ.

If $L(\delta)$ denotes the length of $\Gamma(\delta)$ we have

$$\int_\Omega e^{-2\delta^2/t}\,d\vec{\rho} = \int_0^{\delta_0} e^{-2\delta^2/t}L(\delta)\,d\delta + \text{something less than } |\Omega|\, e^{-2\delta_0{}^2/t}$$

and hence, neglecting an exponentially small term (as well as terms of order t)

$$\int_\Omega e^{-2\delta^2/t}\,d\vec{\rho} \sim \sqrt{t}\int_0^{\delta_0/\sqrt{t}} e^{-2x^2}L(x\sqrt{t})\,dx \sim \sqrt{t}L\int_0^\infty e^{-2x^2}\,dx = \frac{L}{4}\sqrt{2\pi t},$$

where $L=L(0)$ is the length of Γ.

We are finally led to the formula

$$\sum_{n=1}^\infty e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t} - \frac{L}{4}\frac{1}{\sqrt{2\pi t}}, \quad \text{for } t \to 0,$$

and so we can also "hear" the length of the circumference of the drum!

The last asymptotic formula was proved only a few years ago by the Swedish mathematician Ake Pleijel [2] using an entirely different approach.

It is worth remarking that we can now prove that if all the frequencies of a drum are equal to those of a circular drum then the drum must itself be circular. This follows at once from the classical isoperimetric inequality which states that $L^2 \geqq 4\pi|\Omega|$, with equality occurring *only* for a *circle*.

By pitch alone one can thus determine whether a drum is circular or not!

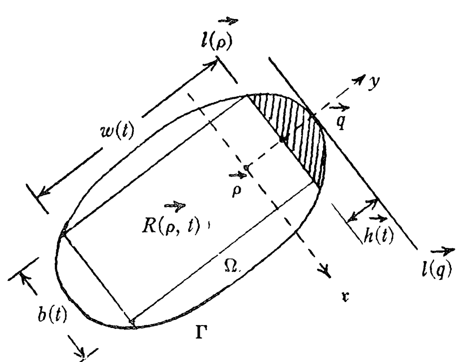


FIG. 5

11. Can the heuristic argument again be made rigorous? Indeed it can. First, we use the inequality

$$P_\Omega(\vec{\rho}\,|\,\vec{\rho};t) \leqq P_{l(\rho)}(\vec{\rho}\,|\,\vec{\rho};t) = \frac{1}{2\pi t} - \frac{1}{2\pi t}e^{-2\delta^2/t},$$

which is simply a refinement of the one used previously, namely, $P_\Omega(\vec{\rho}\,|\,\vec{\rho};\,t) \leqq 1/2\pi t$, and which can be proven the same way.

Next we need a precise lower estimate for $P_\Omega(\vec{\rho}\,|\,\vec{\rho};\,t)$ and this is a little more difficult. We "inscribe" the rectangle $R(\vec{\rho},\,t)$ as shown in Fig. 5, where $h(t)$, the height of the shaded segment, is to be determined a little later.

Let the side of R along the base of the segment be $b(t)$ and the other side be $w(t)$. It should be clear from the picture that the y-axis *bisects* the sides of the rectangle which are parallel to the x-axis.

Now consider $P_R(\vec{\rho}\,|\,\vec{\rho};\,t)$. This notation is perhaps confusing since it suggests that we are dealing with a boundary value problem in which the boundary varies with time. This is not the case. What we have in mind is the following: *fix* t, find $P_{R(t)}(\vec{\rho}\,|\,\vec{r};\,\tau)$ which is defined unambiguously, and finally set $\tau = t$. The result is $P_R(\vec{\rho}\,|\,\vec{r};\,t)$. A convenient expression is

$$P_R(\vec{\rho}\,|\,\vec{\rho};\,t) = \frac{1}{2\pi t}\left\{\sum_{-\infty}^{\infty}\left(\exp\left[-\frac{2b^2}{t}n^2\right] - \exp\left[-\frac{2b^2}{t}\left(n+\frac{1}{2}\right)^2\right]\right)\right\}$$
$$\times\left\{\sum_{-\infty}^{\infty}\left(\exp\left[-\frac{2w^2}{t}n^2\right] - \exp\left[-\frac{2w^2}{t}\left(n+\frac{\bar{\delta}}{w}\right)^2\right]\right)\right\}$$

where $\bar{\delta} = \delta - h(t) = \|\vec{q}-\vec{\rho}\| - h(t)$. Now let $h(t) = \epsilon\sqrt{t}$ and, assuming that $l(\vec{\rho})$ is actually *tangent* to the curve (which for a convex curve will happen with at most a denumerable number of exceptional points $\vec{q}$), we have

$$\lim_{t\to 0}\frac{b(t)}{h(t)} = \lim_{t\to 0}\frac{b(t)}{\epsilon\sqrt{t}} = \infty,$$

and consequently

$$\sum_{-\infty}^{\infty}\left(\exp\left[-\frac{2b^2}{t}n^2\right] - \exp\left[-\frac{2b^2}{t}\left(n+\frac{1}{2}\right)^2\right]\right) = 1 + o(1).$$

This is not quite enough, however, and one needs the stronger estimate

$$\sum_{-\infty}^{\infty}\left(\exp\left[-\frac{2b^2}{t}n^2\right] - \exp\left[-\frac{2b^2}{t}\left(n+\frac{1}{2}\right)^2\right]\right) = 1 + o(\sqrt{t}).$$

This will surely be the case, for example, if the curvature exists at $\vec{q}$, for this would imply that $h(t) \sim b^2(t)$ and the $o(\sqrt{t})$ term above would then be an enormous overestimate. Very mild additional regularity conditions at nearly all points $\vec{q}$ would insure $o(\sqrt{t})$. Without entering into a discussion of these conditions let us simply assume the boundary to be such as to guarantee at least $o(\sqrt{t})$.

Since $w(t)$ remains bounded from below as $t \to 0$, we also have

$$\sum_{-\infty}^{\infty}\left(\exp\left[-\frac{2w^2}{t}n^2\right] - \exp\left[-\frac{2w^2}{t}\left(n+\frac{\bar{\delta}}{w}\right)^2\right]\right)$$
$$= 1 - e^{-2\bar{\delta}^2/t} + \text{exponentially small terms.}$$

We are now almost through. We write (cf. Fig. 6)

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} = \int_{\Omega} P_{\Omega}(\vec{\rho} \mid \vec{\rho}; t)\, d\vec{\rho} > \int_{\Omega(\epsilon\sqrt{t})} P_R(\vec{\rho} \mid \vec{\rho}; t)\, d\vec{\rho}$$

$$= \frac{(1 + o(\sqrt{t}))}{2\pi t} \int_{\Omega(\epsilon\sqrt{t})} (1 - e^{-2\bar{\delta}^2/t} + \text{exponentially small terms})\, d\vec{\rho}.$$

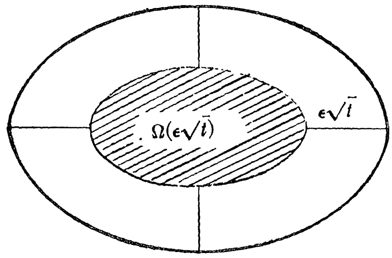


FIG. 6

Except then for exponentially small terms and the factor $1+o(\sqrt{t})$ in front we have the integral

$$\int_{\Omega(\epsilon\sqrt{t})} (1 - e^{-2\bar{\delta}^2/t})\, d\vec{\rho}$$

which, as before, can be seen to be asymptotically

$$|\,\Omega(\epsilon\sqrt{t})\,| - \frac{L}{4}\sqrt{2\pi t},$$

where one neglects terms of order t and exponentially small terms. Since asymptotically $|\Omega(\epsilon\sqrt{t})| \sim |\Omega| - L\epsilon\sqrt{t}$ one can obtain the inequality

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} > \frac{|\,\Omega\,|}{2\pi t} - \frac{(L + \epsilon')}{4} \frac{1}{\sqrt{2\pi t}},$$

where ϵ' is related in a simple way to ϵ. Since ϵ' can be made arbitrarily small, the asymptotic formula

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\,\Omega\,|}{2\pi t} - \frac{L}{4} \frac{1}{\sqrt{2\pi t}}$$

follows.

12. If our overall strategy of attack on the problem is right we should be able to go on and for points very close to a *smooth* boundary replace the boundary locally by suitable circles of curvature.

A result of Pleijel suggests strongly that for a simply connected drum with a smooth boundary (i.e. without corners and with curvature existing at every point) one has

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t} - \frac{L}{4}\frac{1}{\sqrt{2\pi t}} + \frac{1}{6}\,.$$

Unfortunately I am unable to obtain this, for the exasperating reason that I am unable to get a workable expression for $P_\Omega(\vec{\rho}\,|\,\vec{\rho};\,t)$ if Ω is a circle.

Rather than yield to despair over this sad state of affairs let me devote the remainder of the lecture to *polygonal drums*, i.e. drums whose boundaries are polygons. This study will show beyond the shadow of a doubt that the constant term in our asymptotic expansion owes its existence to the overall curvature of the boundary.

13. Before I go on I need an expression for $P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{r};\,t)$ where $S(\theta_0)$ is an infinite wedge of angle θ_0. In other words $P_{S(\theta_0)}$ is the solution of

$$\frac{\partial P}{\partial t} = \frac{1}{2}\,\nabla^2 P$$

subject to the usual initial condition $P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{r};\,t) \to \delta(\vec{\rho}-\vec{r})$, $t\to 0$, and vanishing as $\vec{r}$ approaches a point on either side of the angle θ_0.

This is a very old, very classical, problem and if $\theta_0 = \pi/m$, with m an integer, it can be solved by the familiar method of images. For m not an integer, Sommerfeld invented a method which, so to speak, extends the method of images to a Riemann surface. A little later, in 1899 to be precise, H. S. Carslaw gave a more elementary approach in which $P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{r};\,t)$ is represented by a suitable contour integral. Carslaw transforms the integral into an infinite series of Bessel functions but for our purposes it is best to resist the temptation of Bessel functions and to reduce the integral to a different form. I shall skip the details (though some are quite instructive) and simply reproduce the final result.

Set

$$v(\alpha) = (1/2\pi t) \sum_{\substack{\theta-\alpha-\pi<2k\theta_0\\ <\theta-\alpha+\pi}} \exp\left[-\frac{r^2 - 2r\rho\cos(\theta-\alpha-2k\theta_0)+\rho^2}{2t}\right]$$

$$-\left(\sin\frac{\pi^2}{\theta_0}\right)\frac{\exp\left[-\dfrac{r^2+\rho^2}{2t}\right]}{4\pi\theta_0 t}\int_{-\infty}^{\infty}\frac{\exp\left[-\dfrac{r\rho}{t}\cosh y\right]}{\cosh\left\{\dfrac{\pi}{\theta_0}y+\dfrac{i\pi}{\theta_0}(\theta-\alpha)\right\}-\cos\dfrac{\pi^2}{\theta_0}}\,dy,$$

where the summation $\sum$ is extended over k's satisfying the inequality under the summation sign and $\vec{\rho} = (\rho,\alpha)$, $\vec{r} = (r,\,\theta)$.

Then

$$P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{r};\,t) = v(\alpha) - v(-\alpha).$$

Note that if $\theta_0 = \pi/m$, with m an integer, the complicated integral is out, since the factor in front of it, to wit $\sin \pi^2/\theta_0 = \sin \pi m$, is zero; what remains in the resulting expression for $v(\alpha) - v(-\alpha)$ is a collection of terms easily identifiable with those obtained by the method of images.

Let us now assume that $\pi/2 < \theta_0 < \pi$ and see what $P_S(\vec{\rho}\,|\,\vec{\rho};\,t)$ is in this case. In the expression for $v(\alpha)$ when we set $\theta = \alpha$ the inequality under the $\sum$ sign becomes $-\pi < 2k\theta_0 < \pi$ and only $k = 0$ is allowed. In $v(-\alpha)$ the inequality is $2\alpha - \pi < 2k\theta_0 < 2\alpha + \pi$ and what k's to take depends on α.

We see that:

$$0 < \alpha < \theta_0 - \frac{\pi}{2}, \text{ only } k = 0 \text{ is allowed,}$$

$$\frac{\pi}{2} < \alpha < \theta_0, \text{ only } k = 1 \text{ is allowed,}$$

but for $\theta_0 - \pi/2 < \alpha < \pi/2$ both $k = 0$ and $k = 1$ are allowed. (See Fig. 7.)

Let us now put $\vec{r} = \vec{\rho}$ (so that $\rho = r$) and write down in detail the expressions for $P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{\rho};\,t)$ in the three sectors. For $0 < \alpha < \theta_0 - \pi/2$

$$P_S(\vec{\rho}\,|\,\vec{\rho};\,t) = \frac{1}{2\pi t} - \frac{\exp\left[-\dfrac{r^2}{t}(1 - \cos 2\alpha)\right]}{2\pi t}$$

$$- \left(\sin\frac{\pi^2}{\theta_0}\right)\frac{\exp\left[-\dfrac{r^2}{t}\right]}{4\pi\theta_0 t}\int_{-\infty}^{\infty}\frac{\exp\left[-\dfrac{r^2}{t}\cosh y\right]}{\cosh\dfrac{\pi}{\theta_0}y - \cos\dfrac{\pi^2}{\theta_0}}\,dy$$

$$+ \left(\sin\frac{\pi^2}{\theta_0}\right)\frac{\exp\left[-\dfrac{r^2}{t}\right]}{4\pi\theta_0 t}\int_{-\infty}^{\infty}\frac{\exp\left[-\dfrac{r^2}{t}\cosh y\right]}{\cosh\left\{\dfrac{\pi}{\theta_0}y + 2\pi i\dfrac{\alpha}{\theta_0}\right\} - \cos\dfrac{\pi^2}{\theta_0}}\,dy.$$

For $\pi/2 < \alpha < \theta_0$

$$P_S(\vec{\rho}\,|\,\vec{\rho};\,t) = \frac{1}{2\pi t} - \frac{\exp\left[-\dfrac{r^2}{t}(1 - \cos 2(\theta_0 - \alpha)\right]}{2\pi}$$

$$+ \text{ the same two integrals as above}$$

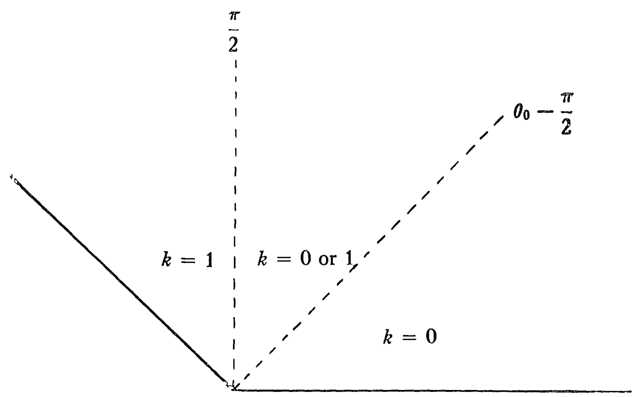


FIG. 7

and finally, for $\theta_0 - \pi/2 < \alpha < \pi/2$

$$P_S(\vec{\rho} \mid \vec{\rho}; t) = \frac{1}{2\pi t} - \frac{\exp\left[-\frac{r^2}{t}(1 - \cos 2\alpha)\right]}{2\pi t} - \frac{\exp\left[-\frac{r^2}{t}(1 - \cos 2(\theta_0 - \alpha))\right]}{2\pi t}$$

$+$ again the same two integrals.

We should recognize $r^2(1-\cos 2\alpha)$ (and $r^2(1-\cos 2(\theta_0-\alpha))$) as being $2\delta^2$ where δ is the distance from $\vec{\rho}$ to a side of the wedge.

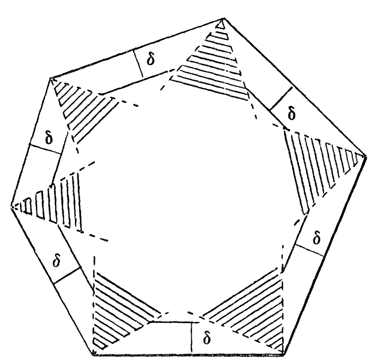


FIG. 8

14. To simplify matters somewhat let me assume that the polygonal drum is convex and that every angle is obtuse.

At each vertex we draw perpendiculars to the sides of the polygon thus obtaining N shaded sectors (where N is the number of sides or vertices of our polygon).

Now let ρ be a point in Ω. Stuff diffusing from $\vec{\rho}$ will either "see" the boundary as a straight line or, if $\vec{\rho}$ is near a vertex, as an infinite wedge.

We may as well say that the boundary will appear to the diffusing particle as the nearest wedge, and that consequently we may replace $P_\Omega(\vec{\rho}\,|\,\vec{\rho};\, t)$ by $P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{\rho};\, t)$, where $S(\theta_0)$ is the wedge nearest to $\vec{\rho}$.

Now, each $P_S(\vec{\rho}\,|\,\vec{\rho};\, t)$ has $1/2\pi t$ as a term and after integration over Ω this gives the principal term $|\Omega|/2\pi t$. Next, each $P_S(\vec{\rho}\,|\,\vec{\rho};\, t)$ contains two complicated looking integrals which have to be integrated over the wedge.

Fortunately, the second of these integrates out to 0, while the first yields, upon integration over $S(\theta_0)$,

$$-\frac{1}{8\pi}\left(\sin\frac{\pi^2}{\theta_0}\right)\int_{-\infty}^{\infty}\frac{dy}{(1+\cosh y)\left(\cosh\dfrac{\pi}{\theta_0}y-\cos\dfrac{\pi^2}{\theta_0}\right)}.$$

This is only the contribution of one wedge; to get the total contribution one must sum over all wedges.

Thus the total contribution is

$$-\frac{1}{8\pi}\sum_{\theta_0}\left(\sin\frac{\pi^2}{\theta_0}\right)\int_{-\infty}^{\infty}\frac{dy}{(1+\cosh y)\left(\cosh\dfrac{\pi}{\theta_0}y-\cos\dfrac{\pi^2}{\theta_0}\right)}.$$

Finally, if $\vec{\rho}$ is in the shaded sector of the wedge $S(\theta_0)$ we get, on integrating over the sector,

$$-\int_{\theta_0-\pi/2}^{\pi/2}d\alpha\int_0^{\infty}\left\{\frac{\exp\left[-\dfrac{r^2}{t}(1-\cos 2\alpha)\right]}{2\pi t}+\frac{\exp\left[-\dfrac{r^2}{t}(1-\cos 2(\theta_0-\alpha))\right]}{2\pi t}\right\}rdr=-\tfrac{1}{2}\,\frac{1}{2\pi}\cot\left(\theta_0-\frac{\pi}{2}\right),$$

and the total contribution from the shaded sectors is $-\frac{1}{2}1/2\pi\sum_{\theta_0}\cot\,(\theta_0-\pi/2)$.

The remaining contribution is easily seen to be

$$-\frac{1}{2\pi t}\int_0^{\infty}\left(L-2\delta\sum_{\theta_0}\cot\left(\theta_0-\frac{\pi}{2}\right)\right)e^{-2\delta^2/t}\,d\delta$$

$$=-\frac{L}{4}\,\frac{1}{\sqrt{2\pi t}}+\tfrac{1}{2}\,\frac{1}{2\pi}\sum_{\theta_0}\cot\left(\theta_0-\frac{\pi}{2}\right).$$

Finally, for a polygonal drum

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t} - \frac{L}{4}\frac{1}{\sqrt{2\pi t}}$$
$$- \frac{1}{8\pi}\sum_{\theta_0}\left(\sin\frac{\pi^2}{\theta_0}\right)\int_{-\infty}^{\infty}\frac{dy}{(1+\cosh y)(\cos \pi/\theta_0\, y - \cos \pi^2/\theta_0)},$$

with the understanding that each θ_0 satisfies the inequality $\pi/2 < \theta_0 < \pi$. If the polygon has N sides, and if we let $N \to \infty$ in such a way that each $\theta_0 \to \pi$, then the constant term approaches

$$+\frac{2\pi}{8\pi}\int_{-\infty}^{\infty}\frac{dy}{(1+\cosh y)^2} = \tfrac{1}{6}\,.$$

This should strengthen our belief that for simply connected smooth drums the constant is universal and equal to $\frac{1}{6}$.

15. What happens for multiply connected drums?

If the drum as well as the holes are polygonal the answer is easily obtained. One only needs $P_{S(\theta_0)}(\vec{\rho}\,|\,\vec{\rho};\, t)$ for θ_0 satisfying the inequality $\pi < \theta_0 < 2\pi$ and this is easily gotten from the general formula quoted above.

Near the holes the diffusing particles will "see" concave wedges but nothing will change in principle.

If we let all polygons approach smooth curves it turns out the constant approaches $(1-r)\frac{1}{6}$, where r is the number of holes. It is thus natural to conjecture that for a *smooth* drum with r *smooth* holes

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t} - \frac{L}{4}\frac{1}{\sqrt{2\pi t}} + (1-r)\,\tfrac{1}{6}\,,$$

and that therefore one can "hear" the connectivity of the drum!

One can, of course, speculate on whether in general one can hear the Euler-Poincaré characteristic and raise all sorts of other interesting questions.

As our study of the polygonal drum shows, the structure of the constant term is quite complex since it combines metric and topological features. Whether these can be properly disentangled remains to be seen.

This is an expanded version of a lecture which was filmed under the auspices of the Committee on Educational Media of the Mathematical Association of America.

References

1. M. Kac, On some connections between probability theory and differential and integral equations, Proc. Second Berkeley Symposium on Mathematical Statistics and Probability, 1957, pp. 189–215.

2. A. Pleijel, A study of certain Green's functions with applications in the theory of vibrating membranes, Arkiv för Matematik, 2 (1954) 553–569.

DERIVATIVES

A. M. BRUCKNER AND J. L. LEONARD, University of California, Santa Barbara

1. Introduction. In recent years there has been a considerable amount of research devoted to questions involving the derivative of a function of one real variable and its generalizations. This activity is due, in part, to the fundamental role played by the derivative in mathematics, and, also, to the difficulty of some of the interesting unsolved problems related to derivatives. It seems appropriate that some of the results of this activity, along with some of the interesting but not-so-well-known earlier results, be brought together and examined in one place. This is one of the purposes of the present expository article.

In deciding which topics to include in this article, we have given preference to ones which can be discussed without first having to develop a great deal of machinery. In addition, we have leaned toward topics in which recent work has been done and for which unsolved problems can be stated.

From the long list of references given at the end of this article, we single out the reference [**192**]. Many of the recent works on derivatives have their origin in this penetrating study by Zahorski.

2. Preliminaries. In this section we present a few of the definitions and concepts which appear in the sequel. To avoid having this discussion become prohibitively long, we restrict ourselves to those notions which appear prominently later on. For other real variable concepts which appear in this article (for example: approximate continuity, F_σ sets, big and little o notation, density conditions) the reader is referred to the texts [**50, 51, 56, 63, 64, 71, 132, 133, 173, 180**].

Throughout this article we shall be concerned with real valued functions of a real variable, usually defined on an interval $[a, b]$. Such a function belongs to Baire class 1 if it is the limit of a sequence of continuous functions. We define the other Baire classes inductively: if α is a countable ordinal, then f is in Baire class α provided it is the limit of a sequence of functions each of which is in a Baire class whose index is less than α. Detailed studies of the Baire classes can be found in [**3, 51, 56, 83, 133, 180**]. We note that our definition is an inclusive one: if f is in Baire class β and $\alpha > \beta$ then f is also in Baire class α. In some studies it is more convenient to have the Baire classes pairwise disjoint. This is the case, for example, in the recent text [**133**]. It is clear that every derivative is a function in Baire class 1.

A property possessed by every derivative of a continuous function is the Darboux property. The function f satisfies the Darboux condition (or f is a Darboux function) on $[a, b]$ provided the set $f[I]$ is connected for every interval $I \subset [a, b]$. This property is often called the *intermediate value property*, because a function has the Darboux property if and only if whenever x_1 and x_2 are points of $[a, b]$ and y is a number between $f(x_1)$ and $f(x_2)$, there is an x_3 between x_1 and x_2 such that $f(x_3) = y$.

Although there are many articles in research journals which deal with Darboux functions, a systematic study of such functions has escaped the standard real variable texts. It is worth noting that the Darboux property is far weaker than the property of continuity. Thus, there exist functions which take on every real value on every perfect set [**60**]. Such a function obviously satisfies the Darboux condition but is nowhere continuous. Another indication of the size of the class of Darboux functions is the fact that *every* function is the limit of a sequence of Darboux functions [**37**]. One more remark: the definition requires that the image (not the graph!) of a connected set be connected. There are Darboux-Baire functions whose graphs are not connected [**84**: p. 82]. However, it is shown in [**85**] that if f is in Baire class 1, then f satisfies the Darboux condition if and only if the graph of f is connected. (See also [**24**; **62**: pp. 289, 290; **84**: p. 81].) (It follows that every derivative has a connected graph.) For further remarks on this subject see [**107**]. The reader interested in Darboux functions is referred to the survey article [**13**].

We end this section by stating for reference the definition of two rather complicated density conditions due to Zahorski [**192**]. We refer to these important conditions in Sections 4, 12, and 14.

DEFINITION. *A nonempty set $E\subset[a, b]$ is said to be an M_4 set provided E is of type F_σ and there exists a sequence $\{F_n\}$ of closed sets and a sequence $\{\eta_n\}$ of numbers, $0<\eta_n<1$, such that $E=\bigcup_{n=1}^{\infty} F_n$ and for each $x\in F_n$ and every $c>0$ there is a number $\epsilon(x, c)>0$ enjoying the following property: for any numbers h and h_1 such that $hh_1>0$, $h/h_1<c$, $|h+h_1|<\epsilon(x, c)$ we have*

$$\frac{m(E\cap(x+h, x+h+h_1))}{|h_1|}>\eta_n.$$

We note that in this definition we require that the numbers η_n can be chosen to be strictly positive. If we relax this requirement to allowing some or all of the η_n to be zero, we arrive at the definition of an M_3 set. The definition of M_3 set can be stated in a different and perhaps simpler form: the set E is an M_3 set provided that if $x\in E$ and $\{I_n\}$ is a sequence of intervals not containing x such that $\{I_n\}\to x$ and $m(I_n\cap E)=0$ for all n, then

$$\lim_{n\to\infty}\frac{mI_n}{\operatorname{dist}(x, I_n)}=0.$$

(The condition M_4 cannot be given in an analogous manner, see Lipiński [**91**].)

3. Continuity of the derivative. The student who has completed a first course in calculus is often not aware of the fact that the derivative of a differentiable function need not be continuous. In a later course he learns that the function f_1 given by $f_1(x)=x^2\sin(1/x)$, $f_1(0)=0$ is differentiable, but f_1' fails to be continuous at the origin. He might never learn, however, just how badly discon-

tinuous a derivative can be. In this section we consider some questions concerning the continuity of derivatives. Our discussion points out, in addition, some of the pathological behavior possible of a derivative.

To show that not every bounded derivative is Riemann integrable, Volterra [**182**] gave an example of a function f_2 whose derivative is bounded but discontinuous on a set of positive (Lebesgue) measure. To construct such a function, he considered a nowhere dense perfect set $P \subset [0, 1]$ of positive measure, and constructed a function which on each interval contiguous to P behaves, roughly, as the function f_1 (above) behaves on $[0, 1]$. This function, f_2, was put together in such a way as to be differentiable on $[0, 1]$ and to produce on all of P the singularity f_1 exhibits at the origin. More precisely, $f_2' = 0$ on P, but f_2' oscillates between -1 and 1 in every neighborhood of an arbitrary point of P. It follows that f_2' must be discontinuous on P. For a precise formulation of such a function see Goffman [**51**: p. 210], Hobson [**63**: pp. 490, 491], or Thielman [**173**: p. 165]. A construction of the type referred to is possible relative to any nowhere dense perfect subset P of an interval I. Such subsets can have measure arbitrarily close to the measure of I, but since $I \sim P$ contains a dense open set, the set P cannot have full measure. It is natural to ask just how large the set of discontinuities of a derivative can be. Does there exist, for example, a derivative which is *everywhere* discontinuous? To see that this question must be answered in the negative, we need only observe that a derivative f' is of Baire class 1 from which it follows that f' must be continuous on a dense set [**133**: p. 143]. We weaken our requirement: is it possible for a derivative to be discontinuous except on a denumerable set? Again the answer is "no." To see this we recall first that the set of points of continuity of *any* function must be a G_δ. Since this set must also be dense, as was seen above, it cannot be denumerable, for a dense G_δ must be nondenumerable. (This fact follows readily from the Baire category theorem.)

We next ask whether or not it is possible for a derivative to be discontinuous on a dense set. To see that this question has an affirmative answer, we use the following approach. We seek a differentiable function whose derivative vanishes on one dense set but is different from zero on another. Such a derivative must, of course, be discontinuous at every point at which it does not vanish. The problem of constructing such derivatives is quite old. In 1887 Köpcke [**77**] claimed to have given an example of a function f possessing the following properties: (a) f has a bounded derivative f' on $[0, 1]$; (b) the set on which f' is positive is dense in $[0, 1]$; and (c) the set on which f' is negative is also dense in $[0, 1]$. Köpcke's original paper had a flaw which he corrected subsequently [**78**, **79**]. There followed a sequence of articles on the subject, culminating in 1915, with a lengthy and penetrating study by Denjoy [**31**]. In this study the author provided a detailed discussion of differentiable functions whose derivatives take on both signs in every interval. He comments on some of the previous works on the subject, including Köpcke's, and gives several methods of constructing such functions. Zalcwasser [**196**] investigated the relative maxima and minima of such functions, obtaining results such as the following: *Let A and B be arbitrary*

nonoverlapping denumerable subsets of $[a, b]$. *Then there exists a differentiable function* f, *having a bounded derivative, such that* A *is the set of strict local maxima of* f *and* B *is the set of strict local minima of* f.

Functions of the Köpcke type are too complicated to discuss here. However, a related kind of function, the so-called function of Pompeiu [**151**], is easier to understand and still exhibits the property that its derivative vanishes on one dense set but is different from zero on another. Again, the derivative of such a function is discontinuous at every point at which it does not vanish. Let us observe that if d is any real number then the function $(x-d)^{1/3}$ has a finite derivative except at d, at which point the derivative is infinite. Let $\{A_n\}$ be any sequence of positive numbers such that $\sum A_n < \infty$, and let $\{d_i\}$ be any denumerable dense subset of $[0, 1]$. Then the series $\sum A_n(x-d_n)^{1/3}$ defines a strictly increasing function f. It can be shown that f has a finite positive derivative at all points for which the differentiated series $\sum \frac{1}{3} A_n(x-d_n)^{-2/3}$ converges, and an infinite derivative otherwise. It can further be shown that the inverse function f^{-1} is a strictly increasing differentiable function and its derivative vanishes on a dense set. A lengthy and detailed study of such functions can be found in Marcus [**106, 113**]. See also [**11**] and [**90**] for answers to some questions raised in [**113**].

Differentiable functions whose derivatives are discontinuous on a preassigned denumerable set $\{d_n\}$ (which may be dense) can be constructed by considering any uniformly convergent series of derivatives, $\sum f_n$, such that the function f_n is discontinuous only at d_n. For example, the function f given by $f(x) = \sum_{n=1}^{\infty} n^{-2} \cos (x-d_n)^{-1}$ is a derivative with discontinuities on the set $\{d_n\}$. (See Halperin [**61**].)

We have seen that the set of discontinuities of a derivative can be dense, but that the set of points of continuity must also be dense and must be nondenumerable. Finally, we ask: what are necessary and sufficient conditions on a set E that it be *the* set of discontinuities of a derivative? It is easy to verify that such a set must be an F_σ of the first category. Conversely, if E is *any* first category F_σ, $E \subset [a, b]$, then E can be expressed as the union of an expanding sequence of nowhere dense closed sets E_n. With each E_n we associate a "Volterra type" function f_n with the property that if $x_0 \in E_n$ then f_n' oscillates between -1 and 1 in each neighborhood of x_0. It is not hard to verify that the function f defined by $f(x) = \sum f_n(x)/3^n$ is differentiable on $[a, b]$ and its derivative is continuous at each point of $\sim E$, but discontinuous at each point of E. Thus we have

THEOREM. *A necessary and sufficient condition that a set* $E \subset [a, b]$ *be the set of discontinuities of a derivative, is that* E *be an* F_σ *of the first category.* (Although we imagine that this theorem is known, we have been unable to find a reference.)

In particular, there are derivatives which are discontinuous a.e. on $[a, b]$. For if to each positive integer n we make correspond a nowhere dense closed subset E_n of $[a, b]$ having measure greater than $b-a-1/n$, then the set E which is the union of the E_n's is a first category F_σ of measure $b-a$. The result follows from the theorem above.

4. An unsolved problem. Many classes of functions can be characterized in terms of what the inverse mapping of a function in the class does to certain open sets. The chart below summarizes some of these characterizations. Let α be any real number and let

$$E_\alpha(f) = \{x: f(x) > \alpha\}, \qquad E^\alpha(f) = \{x: f(x) < \alpha\}.$$

Then f is	*if and only if for all real α, β*
continuous	$E_\alpha(f)$ and $E^\alpha(f)$ are open
Baire class 1	$E_\alpha(f)$ and $E^\alpha(f)$ are sets of type F_σ
Baire class ξ (ξ a countable ordinal)	$E_\alpha(f)$ and $E^\alpha(f)$ are additive Borel class ξ if ξ finite, $\xi+1$ if ξ infinite
in some Baire class	$E_\alpha(f)$ and $E^\alpha(f)$ are Borel sets
upper semi-continuous	$E^\alpha(f)$ is open
lower semi-continuous	$E_\alpha(f)$ is open
measurable	$E^\alpha(f)$ and $E_\alpha(f)$ are measurable
approximately continuous	each $x \in E^\alpha(f) \cap E_\beta(f)$ is a point of density of that set, and that set is an F_σ.

It is natural to ask what the corresponding characterizations are for various classes of derivatives. This question has been studied by Zahorski [**192**]. In this work, he found necessary conditions and also sufficient conditions in terms of the sets $E_\alpha(f)$ and $E^\alpha(f)$ for a function to be a bounded derivative, a finite derivative, or a derivative, possibly infinite, but he was unable to find *characterizations* of these classes of derivatives. (See Section 14 for a more detailed discussion of Zahorski's results.) The question of characterization of these classes is still open.

5. Derivatives a.e. and universal generalized antiderivatives. As we saw in Section 4, the problem of characterizing derivatives in terms of the sets $E^\alpha(f)$ and/or $E_\alpha(f)$ has not yet been resolved. We turn now to the problem of finding such a characterization for the class of functions which have the property of being *almost everywhere* the derivative of a continuous function. As we shall see, this requirement imposes very little restriction on a function.

We first observe that every Lebesgue summable function is almost everywhere the derivative of its integral. The same is true of any function integrable in the sense of Denjoy-Perron. (See Section 10.) There are, however, measurable functions not integrable in either of the above senses. In 1915, Lusin [**97, 99**] published the following theorem, which completely solves the problem of characterizing those functions which are derivatives a.e. of continuous functions: *Every measurable function f (finite a.e.) is almost everywhere the derivative of a continuous function F.* The condition of measurability of f as well as that of a.e. finiteness is obviously necessary as well as sufficient. Thus, *the a.e. finite function f is a.e. the derivative of a continuous function F if and only if f is measurable, or equivalently, if and only if each set of the form $\{x: f(x) < \alpha\}$ or of the form $\{x: f(x) > \alpha\}$ is measurable.* Of course the function f is not the only one

which is a.e. the derivative of F, but any other such function is equivalent to f (i.e., agrees with f a.e.).

Let us see what happens if we weaken the requirement of being a.e. a derivative still further. Let f be an arbitrary function on $[a, b]$ and suppose there exists a sequence $\{h_n\}$ of numbers with $h_n \downarrow 0$ and a continuous function F such that

$$f(x) = \lim_{n\to\infty} \frac{F(x + h_n) - F(x)}{h_n}$$

a.e. on $I = [a, b]$. Then F may be called a *generalized antiderivative* of f. It is clear that such an F may be a generalized antiderivative of many functions not equivalent to f. How many? Marcinkiewicz [**101**] has proved the following remarkable theorem: *There exists a continuous function F which is a generalized antiderivative for every a.e. finite measurable function.* (That is, F is a universal generalized antiderivative.) It was also shown in [**101**] that *most* functions are universal generalized antiderivatives, in that the class of continuous functions which are *not* universal generalized antiderivatives form a set of the first category in $C[a, b]$. A proof of the theorem of Lusin and mention of the theorem of Marcinkiewicz can be found in [**157**: pp. 215–218].

Other results of the above type have been obtained by Sierpiński [**163**] and Eilenberg and Saks [**35**]. One such result [**35**] deals with the generalized antiderivative of arbitrary (not necessarily measurable) functions: *Let f be any function defined on an interval I and let H be any denumerable set of real numbers. Then there exists a continuous function F such that*

$$f(x) = \lim_{n\to\infty} F[(x + h_n) - F(x)]/h_n$$

for every null sequence $\{h_n\}$ from H. (There is no exceptional set here; the result holds for *all* x.)

6. Dini Derivatives. A function defined on an interval $I = [a, b]$ has defined at each point of I four Dini derivatives (except at a and b, at which only two of the Dini derivatives are defined). For example, the upper right Dini derivative of f, $D^+ f$, is defined by

$$D^+f(x) = \limsup_{h\to 0+} \frac{f(x + h) - f(x)}{h},$$

and the other three are defined in an analogous manner. An elementary result is that if one of the Dini derivatives of f is continuous at a point x_0, then f is differentiable at x_0.

In 1915 Denjoy [**30**] proved a theorem relating the four Dini derivatives for continuous functions. This was generalized to measurable functions by Young [**189**] in 1916 and to arbitrary functions by Saks [**156**] in 1924.

THEOREM. *Let f be finite on $[a, b]$. Then, with the possible exception of a null set, $[a, b]$ can be decomposed into four sets:*

A_1, on which f has a finite (ordinary) derivative,

A_2, on which $D^+f=D_-f$ (finite), $D^-f=\infty$, $D_+f=-\infty$,

A_3, on which $D^+f=\infty$, $D_-f=-\infty$, $D^-f=D_+f$ (finite), and

A_4, on which $D^+f=D^-f=\infty$, $D_+f=D_-f=-\infty$.

The theorem is valid if one replaces $[a, b]$ by any set A (not necessarily measurable).

Some immediate consequences of this theorem are the following:

(1) An increasing function is differentiable a.e. (for the sets A_2, A_3, and A_4 are empty in this case);

(2) A function of bounded variation is differentiable a.e. (for such a function is the difference of two increasing functions);

(3) If f is finite on $[a, b]$, then the set on which f' is infinite is a null set. (It is interesting to observe, by way of contrast, that there exist functions which have $D^+f\equiv\infty$, even though f is right-continuous. Here, right-continuity cannot be replaced by continuity. See [**7**: pp. 125, 126; **124**; **167**].)

The Denjoy-Young-Saks theorem has been extended by Garg [**42**], who showed that the exceptional null set also has an image of measure zero. Garg has also considered the set at which the Dini derivatives vanish. Some applications of this extension may be found in [**43, 44, 45**]. For results on Dini derivatives of nowhere monotone functions, the reader is referred to Garg [**46, 47, 48**].

For continuous functions the sets $\{D^+f\neq D^-f\}$ and $\{D_+f\neq D_-f\}$ are small in the sense of category. The following result is due to Neugebauer [**140**]: *If f is continuous then the sets $\{D^+f\neq D^-f\}$ and $\{D_+f\neq D_-f\}$ are of first category. If in addition f is of bounded variation on every closed interval, then these sets are of measure zero as well.* The characteristic function of the rationals shows that continuity cannot be dropped from the first statement; nor can the hypothesis of bounded variation be dropped from the second statement, as is shown by Example III of Denjoy [**30**]. Neugebauer's theorem, as well as certain related results, is a consequence of a result found in [**203**].

Although a Dini derivative does not, in general, satisfy the Darboux condition, some interesting results about the intermediate values taken on by Dini derivatives have been advanced by Morse [**130**]. One such result is the following: *If f is continuous, $-\infty<\lambda<\infty$, if the set $\{x: D^+f(x)\geqq\lambda\}$ is dense and the set $\{x: D^+f(x)<\lambda\}$ is nonempty, then the set $\{x: D^+f(x)=\lambda\}$ has the power of the continuum.*

A derivative (finite or infinite) of a real valued function is always in Baire class 1. The corresponding statement for Dini derivatives is not valid, even for continuous functions. However, if f is in Baire class α, then the four Dini derivatives are in Baire class $\alpha+2$ [**162**] and if f is measurable, then so are its Dini derivatives [**4**]. If f is not measurable, then the same may be true of its Dini derivatives. However, Hájek [**58**] has advanced the surprising result that for any finite function (measurable or not) the extreme *bilateral* derivatives

must be of Baire class 2. (The upper bilateral derivative of a function f is defined by

$$\bar{f}(x) = \limsup_{h\to 0} \frac{f(x+h) - f(x)}{h} \cdot$$

The lower bilateral derivative is defined analogously.) This result is the best possible, for there exists a function f satisfying a Lipschitz condition, with $\bar{f}$ not in Baire class 1 [**168**].

7. Approximate derivatives. In certain instances a function fails to have a derivative at a point x_0, yet the restriction of the function to a set whose complement is very "thin" near x_0 has a derivative at x_0. If one properly interprets "thin" in terms of density, then one arrives at the notion of an approximate derivative.

DEFINITION. *Let f be defined on $[a, b]$, and let $x_0 \in (a, b)$. If there exists a set E such that* (1) $x_0 \in E$, (2) *x_0 is a point of zero density with respect to $\sim E$, and* (3)

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists, for x restricted to E, then this limit is called the approximate derivative of f at x_0 and is written $f'_{ap}(x_0)$. (The obvious modifications are made if $x_0 = a$ or $x_0 = b$.)

The notion of approximate derivative was introduced by Denjoy [**29**] and plays an important role in the theory of the Denjoy-Khintchine integral (see Section 10).

The approximate derivative also arises in connection with certain questions involving the approximation of functions of several real variables. Thus, let f be defined on, say, the unit square S in Euclidean two-space. According to Lusin's Theorem f is measurable if and only if for every $\epsilon > 0$ there exists a function g continuous on S such that $g = f$ except on some set of measure less than ϵ. Suppose we wish to approximate f in this sense by a function which is not only continuous, but also has a continuous total differential. This is possible if and only if f has an approximate total differential almost everywhere [**216**]. That is, in order that f have the property that for every $\epsilon > 0$ there exists a continuously differentiable function g such that $g = f$ except on a set of measure less than ϵ, it is necessary and sufficient that f be approximately differentiable a.e. This last condition is equivalent to the condition that the partial approximate derivatives of f exist a.e. [**213**], see also [**157**: p. 300]. (We note that the corresponding statement for ordinary partial derivatives is false. There is a continuous function of two variables whose partial derivatives exist a.e. but whose total differential exists *nowhere* [**213**: p. 515].)

Approximate derivatives possess some of the properties of ordinary derivatives. Thus, if f is approximately differentiable on $[a, b]$, then f'_{ap} is of Baire

class 1 and possesses the Darboux property (see Khintchine [**73, 74**] and Tolstoff [**174**]). For a unified development of these results, among others, the reader is referred to Goffman and Neugebauer [**53**].

If one allows the approximate derivative to be infinite, then the situation is a little different. Zahorski [**191**] has given an example of a function f which has at every point an approximate derivative (finite or infinite), but f'_{ap} is not of Baire class 1, nor does it satisfy the Darboux condition. In the same article, however, he shows that if f has at each point an approximate derivative (finite or infinite), then f'_{ap} is of Baire class 2. In addition, f must be of Baire class 2 as well. In case f is approximately continuous and has at each point a finite or infinite approximate derivative, then f'_{ap} (as well as f) must be of Baire class 1. (See Tolstoff, [**174**].) In addition, f'_{ap} must satisfy the Darboux condition in this case [**82**]. Some additional results involving the Baire class of approximate derivatives can be found in Krzyzewski [**81**] and Matysiak [**119**].

It is of interest to note that while the set of discontinuities of a function having everywhere a derivative (possibly infinite) must be denumerable, Lipiński [**92**] has shown that the set of points of approximate discontinuity of a function having everywhere an approximate derivative (possibly infinite) can be nondenumerable, although the set must have zero measure and be of the first category.

Under certain conditions a point of approximate differentiability is actually a point of differentiability. Thus if f is monotonic, f is differentiable wherever f is approximately differentiable [**73, 74**]. Khintchine has also shown that if an approximate derivative (possibly infinite) is dominated by an ordinary derivative, then this approximate derivative is in fact an ordinary derivative. This result has been used by Tolstoff to prove that if f is approximately continuous and has at each point a finite or infinite approximate derivative f'_{ap}, then except possibly on a nowhere dense set, f'_{ap} is the ordinary derivative of f. ([**175**]; see also [**53**].)

In 1916 Denjoy [**33**] noted the following property of a (finite) derivative: *If $\alpha<\beta$ then the set $E_{\alpha\beta}=\{x\colon \alpha<f'(x)<\beta\}$ is either empty or has positive measure.* This result was extended by Clarkson [**26**] to derivatives (which might be infinite) of continuous functions. A more detailed description of the sets $E_{\alpha\beta}$ was advanced by Hsiang [**65**]. Finally, in 1962, Marcus [**108**] showed that the corresponding results are valid if one replaces "derivative" by "approximate derivative" in the hypothesis and conclusion of Clarkson's theorem. In case one allows the approximate derivative to be infinite at some points, additional assumptions are necessary. As mentioned in Section 4, Zahorski [**192**] has obtained results concerning the structure of sets of the form $\{x\colon f'(x)<\beta\}$. Corresponding results for approximate derivatives and Peano derivatives have been advanced by Weil [**185**] and Kulbacka [**82**].

Whether or not a function is approximately differentiable, it always has four extreme unilateral approximate derivates. For a detailed study of these, the reader is referred to Jeffery [**71**]. We mention the interesting fact, which may be

found on pages 198–199 of [**71**], that the theorem of Denjoy-Saks-Young for Dini derivatives (see Section 6 above) has a virtually identical analogue. If f happens to be measurable, then the sets which correspond to A_2 and A_3 (of Section 6) are null sets. For results on nonmeasurable functions, see Chow [**25**].

It is of interest to note that for functions of *several* variables the extreme unilateral partial approximate derivatives of a function f reflect the measurability properties of f, whereas the partial Dini derivatives do not. Thus if f is a Lebesgue (Borel) measurable real valued function of several variables, the same is true of its extreme unilateral partial approximate derivatives. On the other hand, there are Lebesgue (Borel) measurable functions of two variables whose partial Dini derivatives are not Lebesgue (resp. Borel) measurable. It is true, however, that if f is continuous (Borel measurable), then its partial Dini derivatives are Borel measurable (resp. Lebesgue measurable). (For functions of one real variable, the Dini derivatives as well as the extreme unilateral approximate derivatives inherit the Lebesgue or Borel measurability of the primitive function.) For results of this sort see [**83**: p. 421], [**157**: pp. 113, 171, 299], and [**209**].

We conclude by mentioning that the word "thin" mentioned in the introductory paragraph can be interpreted in other ways, giving rise to different sorts of derivatives. Thus, for example, S. Marcus [**111, 112**] has interpreted "thin" in terms of category (rather than measure) and arrived at the notion of a qualitative derivative. The notion of "preponderant" derivative, due to Denjoy [**29**], is related to the notion of approximate derivative but the complements need not be quite so "thin" for a preponderant derivative to exist as for an approximate derivative to exist. For a fuller discussion of this matter in a slightly broader context, consult Section 8.

8. Other generalizations of the derivative. It is scarcely surprising that such a fundamental concept as the derivative has received generalization in a number of different directions for various special purposes. Many generalizations are arrived at by weakening the sense in which the limit of the difference quotient $[f(x+h)-f(x)]/h$ is obtained, although other avenues of definition are sometimes used. Usually the existence of the generalized derivative together with some regularity condition implies the existence of the ordinary derivative, and the restrictiveness of this regularity condition can be used as an index of the degree of generalization obtained. Where the ordinary derivative exists, it is equal to the generalized derivative. We shall give the definitions of various generalizations and discuss briefly some of the more important among them.

The Dini derivates, discussed in Section 6, represent the first generalization of the ordinary derivative, in that we do not restrict ourselves to the limit of the difference quotient, which limit may fail to exist, but rather consider the one-sided limit inferior and limit superior. In this way we are assured of the existence of these derivatives at each two-sided limit point of the domain of the function. We need only be given the continuity of one Dini derivate at a point to conclude the existence of the ordinary derivative at that point.

The approximate derivative, discussed in Section 7, is a natural generalization of the ordinary derivative in which the limit of the difference quotient is taken in the metric sense of the approximate limit. As pointed out in Section 7, an approximately continuous function with an approximate derivative everywhere in an interval $[a, b]$ possesses an ordinary derivative on a set of intervals which is dense in $[a, b]$ (see [**174**]).

Weakening the density requirements for the existence of a limit in the definition of approximate derivative (see Section 7) to *the set E has mean density greater than* $1/2$ *on all sufficiently small intervals including* x_0, we obtain the *preponderant* derivates and derivative of Denjoy [**29**]. Replacing metric considerations with the concept of category, one arrives at the *approximate qualitative* derivative defined by S. Marcus [**111, 112**], where the upper qualitative limit of f at x_0 is defined as $\inf \{y: \{x: f(x) > y\}$ is first category at $x_0\}$, the lower qualitative limit is similarly defined, and these limiting operations are applied to the difference quotient to yield approximate qualitative derivates, which share many of the properties of the Dini derivates.

In taking the limit of the difference quotient, we may restrict ourselves to considering only values of x_0+h which belong to a given set E, which set has x_0 as a limit point. This will give us the *derivative of f relative to the set E*. Many of the theorems found in Saks [**157**] hold for this form of the derivative. Similar in concept is the *congruent derivative* of Sindalovskiĭ [**164, 165, 212**], in which the values of h used in forming the difference quotient are restricted to belong to a set Q, which has 0 as a limit point, but where this difference quotient is defined for every $x \in [a, b]$, always using the same set Q. The idea of passing to the limit while neglecting values obtained on "negligible" sets belonging to a particular family has been advanced by Császár [**27**].

Changing the form of the difference quotient gives rise to many generalizations of the derivative. The most common is the symmetric derivative (also called the Riemann derivative), defined by

$$f^{[\prime]}(x) = \lim_{h\to 0} \frac{1}{2h}\,[f(x+h) - f(x-h)].$$

This derivative has the virtue of not involving the behavior of f at the point x itself. It is widely used in the theory of trigonometric series. (See, for example, [**198, 199**].) The existence of the symmetric derivative at all points of a set E implies the existence of the ordinary derivative a.e. in E [**73**]. The symmetric derivative is nicely arrived at through decomposing the function f at x_0 into its *even* and *odd* parts:

$$\phi_{x_0}(t) = \tfrac{1}{2}[f(x_0+t) + f(x_0-t)],$$
$$\psi_{x_0}(t) = \tfrac{1}{2}[f(x_0+t) - f(x_0-t)].$$

The differentiability of the odd part at $t=0$ is then equivalent to the existence of the symmetric derivative, while the differentiability of the even part is equiv-

alent to the property of *smoothness* [**197**, **139**] (see Section 9). The higher Riemann derivatives, given by

$$f^{[n]}(x_0) = \lim_{h \to 0} \frac{1}{h^n} \sum_{k=0}^{n} (-1)^k f\left(x + h\left(\frac{n}{2} - k\right)\right),$$

extend the symmetric derivative [**18**, **20**, **72**, **199**]. The second Riemann derivative is often called the Schwarz derivative.

The difference quotient may be varied in other directions. The simplest variant is

$$f^*(x_0) = \lim_{\substack{x_1, x_2 \to x_0 \\ x_1 \neq x_2}} \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

This definition was considered by Peano [**146**], who felt that it portrayed the concept of the derivative used in the physical sciences more closely than does the usual definition, since f^* is always continuous, and coincides with f' whenever f' is continuous. This definition has been recently reconsidered by Esser and Shisha [**36**]. Another variation is Sindalovskiĭ's derivative [**166**],

$$\lim_{h \to 0} \frac{f(x - \phi(h)) - f(x - \phi(h) - h)}{h},$$

where ϕ is an arbitrary function, defined in a neighborhood of the origin, which approaches 0 with h. Murav'ev [**131**] dealt with the Gateaux derivative,

$$\lim_{h \to 0} \frac{f(x + h\alpha(x)) - f(x)}{h},$$

where f is differentiable (in the ordinary sense) on $[a, b]$, and $\alpha(x)$ is any bounded function defined on $[a, b]$. The linear function in the denominator of the difference quotient may be exchanged for an arbitrary function $g(x)$, yielding the *derivative with respect to* g given by

$$\lim_{x \to x_0} \frac{f(x_0) - f(x)}{g(x_0) - g(x)} \qquad [\mathbf{88}].$$

The physical sciences, in particular thermodynamics, led Borel to define a *mean derivative* [**10**],

$$f_B'(x) = \lim_{h \to 0} \lim_{\epsilon \to 0} \frac{1}{h} \int_{\epsilon}^{h} \frac{f(x + t) - f(x)}{t}\, dt.$$

Sargent [**158**] extended this definition to parallel the Dini derivates, and Marcinkiewicz and Zygmund [**103**] extended his results to a "smooth Borel derivative."

In a masterly paper, Khintchine [**73**] considered various candidates for a "generalized derivative," establishing their properties and constructing examples to illustrate a hierarchy of generality. The symmetric derivative was discarded since its existence, save on a null set, implies the existence of the ordinary derivative a.e. The Borel derivative generalizes the symmetric derivative and lacks this "flaw," but is itself generalized by the approximate derivative. The approximate derivative, however, is generalized by *la dérivée générale* of f, which is a function f_g', defined only a.e., such that for any $\epsilon > 0$,

$$m\left\{x: \left|\frac{f(x+h)-f(x)}{h} - f_g'(x)\right| > \epsilon\right\}$$

tends to 0 with h. The existence of f_{ap}' a.e. on an interval implies the existence of f_g' on the interval, and $f_{ap}' = f_g'$ a.e., while Khintchine constructed a function f such that f_g' exists on $[0, 1]$, and f_{ap}' only on a null set. However, f_g' bows to *la dérivée généralisée*, f_G', defined even less uniquely, which is any function with the property that, for some sequence $\{h_n\}$ decreasing to 0, we have $[f(x+h_n)-f(x)]/h_n \to f_G'$ a.e. This treatment reflects the generalized antiderivative of Section 5. Khintchine closed by constructing a continuous function which fails to have even a *dérivée généralisée.*

A completely different approach, which yields a generalization of ordinary derivatives of order greater than one, is given by polynomial approximation to a function. If $f(x_0+h)$ can be expressed as

$$f(x_0+h) = f(x_0) + hf_1(x_0) + \frac{h^2}{2!}f_2(x_0) + \cdots + \frac{h^n}{n!}f_n(x_0) + o(h^n),$$

then the f_i's are referred to as the ith Peano derivatives [**145**] (referred to by Denjoy [**32**] as differential coefficients, and sometimes called de la Vallée Poussin derivatives [**181**], although this latter term is also used otherwise [**199**]). The nth Peano derivative f_n always equals the nth ordinary derivative when the latter exists. Oliver [**143**] studied the exact nth Peano derivative, which is one which exists at every point of an interval, and showed that it is of Baire class 1, enjoys the Darboux property and the Denjoy property of $E_{\alpha\beta}$, and coincides with the ordinary nth derivative on a dense open set. A side condition for the existence everywhere of the nth ordinary derivative is that the nth Peano derivative be bounded either above or below. If the nth Riemann derivative exists everywhere on a set E of positive measure, then the nth Peano derivative exists a.e. on E [**103**].

Generalization of the Peano derivative leads to the L^p derivative: if $f \in L^p$, $1 \leqq p \leqq \infty$, in some neighborhood of x_0, and if a polynomial

$$P(t) = \sum_{i=0}^{n} \frac{f_i}{i!} t^i$$

exists such that

$$\left[\frac{1}{2h}\int_{-h}^{h}|f(x_0+t)-P(t)|^p\,dt\right]^{1/p}=o(h^n),$$

then f is said to be differentiable of order n at x_0 in L^p, and f_i is the ith L^p derivative. This generalization was introduced by Calderón and Zygmund [**21, 22**] because the property of differentiability in L^p at a point is preserved under various integral transformations, and was applied to solving partial differential equations. The L^p derivative has been used recently to establish the differentiability a.e. of functions [**136, 138, 169**] (see Section 9). Weiss [**186**] considered the symmetric kth derivative in L^p, and generalized the result that the existence of the kth symmetric derivative implies the existence of the kth (Peano) derivative a.e. Higher dimensions are considered in [**215**].

The nth Taylor derivative also arises out of considerations of polynomial approximation. It is defined [**19, 20**] as

$$f^{\{n\}}(x)=\lim_{h\to 0}\frac{n!}{h^n}\left[f(x+h)-\sum_{k=0}^{n-1}\frac{h^k}{k!}f^{(k)}(x)\right],$$

where $f^{(k)}$ is the ordinary kth derivative. Butzer [**19**] and Görlich and Nessel [**55**] studied the relationships between the Riemann, Peano, Taylor, and ordinary nth derivatives (listed in order of decreasing generality), where convergence was considered in both the usual metric and in the L^p norm.

For those interested in further generalizations of the derivative, we mention the fluents [**126**] and multiderivative [**125, 127**] of Menger, the Hölder and Cesaro derivatives of order n [**152**], the fractional derivative of Kuttner [**86**], Minetti's right oscillatory derivative [**128**], O'Neill's work on generalized derivates [**144**], and Shukla's on nonsymmetric differentiability [**160, 161**].

The extension of the concept of derivative to spaces other than the real line gives rise to a vast literature, which we will not explore. Introductions to this field are provided by Bögel's paper on higher-dimensional differentiation [**9**], Fréchet's study [**39**] of various definitions of differentiability in the plane, and the article of Rinehart and Wilson on differentiation in algebras [**153**].

9. Points of differentiability. Since the existence of the derivative of a function f at a point implies a certain degree of good behavior of f at that point, it is most natural to study the set of such points of differentiability. It is especially desirable to find under what conditions this set is large, and it is also of interest to know something about the size of the image of this set under f. We discuss these questions in this section.

If f' is defined and finite at x then x is a point of differentiability of f; if f' is defined (possibly infinite) at x, then we call x a point of extended differentiability. We let D represent the set of points of differentiability of f, and D^* the points of extended differentiability, and also let $N=\sim D$, $N^*=\sim D^*$.

The first result encountered is that a function which is monotonic on an interval has a finite derivative almost everywhere on that interval. The conclu-

sion carries over easily to functions of bounded variation (Lebesgue's Theorem [**157**: p. 223]), and to functions which are VBG_*. Indeed, in these cases, not only is $mN=0$, but $mf[N^*]=0$. ([**157**: p. 230], due originally to Denjoy and Lusin). A condition, phrased in terms of a more restrictive notion of absolute continuity, which is both necessary and sufficient for f to be differentiable almost everywhere, under the restriction that f is continuous, has been presented by Pettineo [**149, 150**].

The property of smoothness (see Section 8) and the related property Λ are of interest in investigating the differentiability of a function [**197, 138, 170**]. The function f is smooth at x_0 if

$$\Delta_x(h) = f(x+h) + f(x-h) - 2f(x) = o(h);$$

f satisfies condition Λ if this $\Delta_x(h) = O(h)$.

A continuous smooth function f on (a, b) has the property that the set D of its points of differentiability has the power of the continuum in every sub-interval of (a, b) [**197**]. Furthermore, f' satisfies the Darboux condition on D. (See [**197**].) If the continuity of f is replaced by measurability, the result on differentiability still holds, but in order to conclude the Darboux property of f' we need to know that D is "small" in the sense that $m(D\cap I) < m(I)$ for each interval $I \subset (a, b)$ [**139**].

If f satisfies condition Λ, then we may state a necessary and sufficient condition for the differentiability of f a.e.: *A measurable function f satisfying condition Λ at each point of a measurable set E is differentiable* a.e. *on E if and only if for almost every $x \in E$ there is an η_x such that $h^{-1}[\Delta_x(h)]^2$ is summable over* $(0, \eta_x)$. The necessity is due to Marcinkiewicz [**102**], and the sufficiency to Stein and Zygmund [**169**]. The summability of $h^{-1}[\Delta_x(h)]^2$ a.e. is equivalent to the existence of the L^2 derivative a.e. [**169**].

The condition Λ is dispensed with in a similar result due to Neugebauer [**136**]: *A measurable function f is equivalent to a function differentiable a.e. on a measurable set E if and only if for almost every $x \in E$ there is an $\eta_x > 0$ such that $[\Delta_x(h)]^2/h^3\phi(h^{-1}\Delta_x(h))$ is summable over* $(0, \eta_x)$, where the function ϕ is given by $\phi(x) = 1 - |x|$ in $(-1, 1)$, $\phi(x) = 0$ elsewhere.

Properties which are relevant to a discussion of differentiability are Banach's conditions (T_1) and (T_2) and Lusin's condition (N). For f defined on $I=[a, b]$ and $y \in f[I]$, we call the set $\{x: f(x) = y\}$ a level set of f. We say that f satisfies condition (T_1) if, for almost all $y \in f[I]$, the level sets are finite, and (T_2) if the level sets are at most denumerable. The condition (N) is satisfied if $B \subset I$ and $mB = 0$ imply that $mf[B] = 0$ also.

Marchaud [**100**] showed that if each level set of a continuous function f is finite, then f is differentiable almost everywhere. Iosifescu [**68**] has given a direct demonstration of this theorem, and has extended the result to discontinuous functions for which the set of points of nonmonotonicity (i.e., points having no neighborhood on which the function is monotonic) has measure zero. If the finiteness of the level sets is extended to denumerability, the result is al-

most totally lost, for Iosifescu has given a construction of such a function f_ϵ, for which $mD < \epsilon$, where ϵ is any arbitrarily preassigned positive number [**67**].

If f is continuous and satisfies property (N), then a theorem of Banach's [**157**: p. 286] assures us that D has positive measure. Weakening the hypothesis by replacing (N) with (T_1), we can conclude only that N^* has an image of measure zero (a property that characterizes continuous functions which are (T_1) [**157**: p. 278]). Still further weakening the hypothesis to (T_2), we can conclude only that D^* is nondenumerable, but can say nothing about its measure.

Finally, we know that it is possible for a continuous function to be so badly behaved as to be nowhere differentiable. It may be that D is empty but D^* is not empty, as in Cellerier's example [**23**], or we may even have D^* empty, as in Weierstrass' function [**184**], which does, however, admit one-sided derivatives on a dense set. Even this last remnant of good behavior can be removed, as in Besicovitch's example [**6, 147**], which at no point has even a unilateral derivative (even infinite). These are all discussed in Jeffery [**71**]. Functions such as Besicovitch's are "much rarer" than those of Weierstrass' example, in the sense that the former constitute a first category set in $C[a,b]$ while the latter form the complement of a first category set [**5, 123, 155**].

We pass from the study of the size of D and N to considerations of the structure of these sets. Through use of the concept of convergence classes [**62**: p. 309] one can prove that D is an $F_{\sigma\delta}$. The same is true of the set of points of left-differentiability or points of right-differentiability. It is not the case, however, that each $G_{\delta\sigma}$ is the set N for some function f. Zahorski [**190, 193**] showed for continuous functions that the set N is the union of a G_δ with a null $G_{\delta\sigma}$ and that any set of this form is the set of points of nondifferentiability for some continuous function. Exactly the same statement is true of N^*. For a function of bounded variation, the G_δ is dropped from the theorem. Brudno [**17**] extended the results to arbitrary functions, with exactly the same conditions holding. Zahorski's proof has been simplified by Piranian [**211**].

Since the distinction between having a derivative (possibly infinite) and having a finite derivative is so often critical (see, for example, Section 10 on inversion of derivatives), it is of interest to know just where a derivative may take on infinite values. We know from our discussion of Denjoy's theorem on Dini derivatives that the set $\{x: f'(x) \text{ is infinite}\}$ has measure zero. Conversely, for any set E of measure zero, there is a simple construction of a continuous, increasing function f_E with $f_E' = +\infty$ on E (see [**132**: p. 214]). Jarník [**70**] gave a construction of a continuous function with an infinite derivative on an arbitrarily given null G_δ and with finite Dini derivatives elsewhere, and Zahorski [**195**] improved this result to present an everywhere differentiable (in the extended sense) function with this property. For other results of this nature see Bojarski [**200**], Lipiński [**208**], Marcus [**116**], and Piranian [**210**].

The most complete result in this direction is due to Tzodiks [**177, 178, 179**]: *For a finite function f, necessary and sufficient conditions for the sets E_1 and E_2 to be sets where $f' = +\infty$ and $f' = -\infty$ respectively are*: (1) *E_1 and E_2 be $F_{\sigma\delta}$'s with*

measure zero, and (2) *there exist disjoint* F_σ*'s* H_1 *and* H_2, *with* $E_1 \subset H_1$, $E_2 \subset H_2$. Other results concerning infinite derivatives are given in Filipczak [**38, 204**], Garg [**49**], Kronrod [**80**], Landis [**87**], Marcus [**109, 116**], and Marczewski [**118**].

10. Inversion of derivatives. We shall be concerned in this section with that half of the fundamental theorem of calculus which, roughly, recaptures a function from its derivative. The form which this theorem usually takes in elementary calculus is: *Let f be continuously differentiable on* $[a, b]$. *Then*

$$f(b) - f(a) = \int_a^b f'(x)\,dx, \tag{*}$$

the integral being taken in the sense of Riemann. The requirement that f' be continuous is usually weakened in a course in advanced calculus to the requirement that f' be Riemann integrable. The example of Volterra cited in Section 3 shows that this latter restriction cannot be weakened to insisting merely that f' be bounded. Now a desirable property of an integral is that the fundamental equation (*) hold for *any* derivative f', irrespective of whether or not f' is continuous or bounded. The above consideration shows that the Riemann integral does not have this property, even for bounded derivatives. The Lebesgue integral does a little better. The relevant theorem for Lebesgue integrals asserts that (*) holds whenever f' is summable. In particular, (*) holds for Lebesgue integrals whenever f' is bounded. If f' is not bounded, then f' might fail to be summable. The function $f(x) = x^2 \sin x^{-2}$, $f(0) = 0$ furnishes an example of a differentiable function on $[0, 1]$ whose derivative is not summable over any interval containing the origin. The difficulty, of course, lies in the fact that $\int |f'| = \infty$ over any such interval. It is of interest to note that even derivatives which are "tied down" by vanishing on a dense set of points (see Section 3) can be so large elsewhere that f' fails to be summable. (See [**11**] and [**90**].) We have seen that equation (*) is not valid for Lebesgue integrals in general.

Perron [**148**] and Denjoy [**28, 34**] independently defined integrals, both more general than the integral of Lebesgue's, which completely solved the problem of recapturing a function from its (finite) derivative; more precisely, of integrating arbitrary derivatives so that (*) holds. Although the methods of Denjoy and Perron were entirely different in approach, Hake [**59**], Alexandroff [**1**, 2], and Looman [**96**] proved that these two integrals were entirely equivalent; i.e., if a function is integrable in one of the two senses, it is integrable in the other, and the two integrals are equal. Thus, this integral is usually called the Denjoy-Perron integral.

In 1916 Khintchine [**75, 76**] modified the Denjoy construction to give rise to a more general integral, now referred to as the Denjoy-Khintchine integral, which integrated arbitrary approximate derivatives of continuous functions. *Descriptive* definitions of the Denjoy-Perron and Denjoy-Khintchine integrals

were advanced by Lusin [**98**]. For a development of the Perron integral and both the constructive and descriptive definitions of the Denjoy-Perron and Denjoy-Khintchine integrals, see Saks [**157**]. An elegant development of the Perron and the Denjoy integrals, along with a proof of their equivalence, can be found in Natanson [**133**; Chap. 16]. A detailed discussion of how a function may be recaptured from its derivative in a countable number of steps is given in Jeffery [**71**].

For purposes of comparison we state the descriptive definitions of the Lebesgue, Denjoy-Perron, and Denjoy-Khintchine integrals. We begin with the definitions of four generalizations of the notion of absolute continuity of a function defined on an interval $[a, b]$: Let F be continuous on $[a, b]$ and let $E \subset [a, b]$. Then F is called $AC(AC_*)$ on E provided that for any $\epsilon > 0$ there exists a $\delta > 0$ such that if $\{[a_k, b_k]\}$ is any sequence of nonoverlapping intervals with endpoints in E and with $\sum(b_k - a_k) < \delta$, then $\sum|f(b_k) - f(a_k)| < \epsilon$ ($\sum\omega_k < \epsilon$, where ω_k denotes the oscillation of f on $[a_k, b_k]$). If $[a, b] = \bigcup E_k$, such that F is $AC(AC_*)$ on each set E_k, then F is called $ACG(ACG_*)$ on $[a, b]$.

Descriptive definition of the Lebesgue integral: The function F is called the Lebesgue integral of a function f provided that:

(a) F is absolutely continuous on $[a, b]$, and

(b) $F' = f$ a.e.

Descriptive definition of the Denjoy-Perron integral: The function F is called the Denjoy-Perron integral of a function f provided that:

(a) F is ACG_* on $[a, b]$, and

(b) $F' = f$ a.e.

Descriptive definition of the Denjoy-Khintchine integral: The function F is called the Denjoy-Khintchine integral of a function f provided:

(a) F is ACG on $[a, b]$, and

(b) $F'_{ap} = f$ a.e.

For a detailed development of the relevant concepts, the reader is referred to Saks [**157**].

We conclude by observing that the derivatives considered in this section are taken to be finite. This requirement cannot be entirely deleted, for there exist two continuous functions F and G, such that $F' \equiv G'$, yet $F - G$ is not constant. These derivatives are equal to $+\infty$ on the Cantor set, and finite elsewhere. The difference $F - G$ is the Cantor function. The first to notice the existence of two such functions was Hahn [**57**]. See, also, Ruziewicz [**154**], and Saks [**157**: pp. 205, 206].

11. Stationary sets and determining sets. A standard theorem of elementary calculus asserts that if the derivative of a differentiable function vanishes on an interval, then the function is constant. One might ask the question: "On how large a set must the derivative be known to vanish, before it is known to vanish identically?" This leads us to the notion of a stationary set for a class of functions.

DEFINITION. *Let $\mathcal{C}$ be a collection of functions defined on $[a, b]$. A subset E of $[a, b]$ with the property that whenever $f \in \mathcal{C}$ is constant on E, then f must be constant on $[a, b]$, is said to be a stationary set for $\mathcal{C}$.*

For example, if $\mathcal{C}$ consists of the continuous functions on $[a, b]$, then the stationary sets for $\mathcal{C}$ are the dense sets, while if $\mathcal{C}$ consists of the analytic functions, then the stationary sets for $\mathcal{C}$ are those which contain at least one limit point.

If the collection of functions $\mathcal{C}$ is closed under the operation of subtraction, then every stationary set for $\mathcal{C}$ is also a *determining set* for $\mathcal{C}$. That is, two members of $\mathcal{C}$ which agree on this set must agree on all of $[a, b]$.

In recent years, the stationary sets and the determining sets for various classes of derivatives, as well as certain related classes of functions, have been characterized. See Boboc and Marcus [**8**], Bruckner [**12**], Bruckner and Leonard [**16**], Goffman and Neugebauer [**52**], Marcus [**104, 105, 110, 114, 115, 117**], Neugebauer [**134**], Sunyer Balaguer [**171, 172**]. The results of those investigations which bear directly on our subject are tabulated in the chart below, which lists the characterizations of stationary sets and determining sets for various classes of functions defined on $[a, b]$. If A is a set then $m_i(A)$ denotes its Lebesgue inner measure and card (A) its cardinality.

12. Intervals of constancy. An interesting function encountered by students of a course in real variables is the Cantor function f. This function is defined on $[0, 1]$ with $f(0) = 0$, $f(1) = 1$. It has the property that although it is continuous and nondecreasing on $[0, 1]$, it is constant on every interval contiguous to the Cantor set P. Thus $f' = 0$ except on P. It can be shown, however, that f fails to have a derivative, finite or infinite, on a non-denumerable set. A natural question to ask is whether one can in some way "smoothe" the Cantor function to arrive at a differentiable function f such that $f(0) = 0$, $f(1) = 1$, and $f' = 0$ on $\sim P$. The answer is in the negative, in view of the following result due to Zahorski [**192**: p. 21]: *If a continuous nonconstant function f of bounded variation has almost everywhere a vanishing derivative, then f fails to be differentiable on an uncountable set.*

This theorem does not eliminate the possibility that a nonconstant function be differentiable on an interval, yet constant on each interval of a set of intervals whose union is dense in $[0, 1]$. Such functions have actually been constructed; see Zahorski [**194**]. In fact, the following statement is valid [**192**; p. 43]: *A necessary and sufficient condition that E be the set of zeroes of a bounded derivative is that $\sim E$ be an M_4 set.* Using this theorem one can prove the following [**15**]: *Let G be an open dense subset of $[a, b]$ and let $P = \sim G$. A necessary and sufficient condition that there be a differentiable function f defined on $[a, b]$ such that f is constant on each component interval of G, but not constant on any open interval containing points of P, is that the intersection of P with any arbitrary open interval is either empty or has positive measure.*

Class of functions	E is a stationary set if and only if	E is a determining set if and only if
I. *Derivatives*		
Derivatives (possibly infinite)	$E = [a, b]$	$E = [a, b]$
Derivatives (possibly infinite) of continuous functions	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
Finite derivatives	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
Bounded derivatives	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
Riemann integrable derivatives	E is dense	E is dense
Bounded semicontinuous derivatives	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
II. *Approximate derivatives*		
Approximate derivatives (possibly infinite)	$E = [a, b]$	$E = [a, b]$
Approximate derivatives (possibly infinite) of Darboux functions	(See Note Below)	$E = [a, b]$
Approximate derivatives (possibly infinite) of approximately continuous functions	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
Approximate derivatives (possibly infinite) of continuous functions	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
Finite approximate derivatives	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$
III. *Dini derivatives*		
Dini derivatives of Darboux Baire functions	E meets every perfect set	$E = [a, b]$
Dini derivatives of continuous functions	E meets every perfect set	E meets every perfect set
Finite Dini derivatives of continuous functions	E meets every perfect set	E meets every perfect set
IV. *Darboux functions*		
Darboux functions	card $(\sim E) < c$	$E = [a, b]$
Measurable Darboux functions	E meets every uncountable measurable set	$E = [a, b]$
Darboux Baire functions	E meets every perfect set	$E = [a, b]$
Darboux Baire class 1 functions	E meets every perfect set	$E = [a, b]$
Lower semicontinuous Darboux functions	E meets every perfect set	E meets every perfect set
Approximately continuous lower semicontinuous functions	$m_i(\sim E) = 0$	$m_i(\sim E) = 0$

Note. A necessary condition for E to be a stationary set for the class of approximate derivatives (possibly infinite) of Darboux functions is that $m_i(\sim E)=0$; a sufficient condition is that E meet every perfect set [12].

We also note that the stationary sets and the determining sets for both the class of approximately derivable functions and the uniform closure of this class are the sets which are dense in the interval $[a, b]$ [**134**].

13. Monotonicity. According to a theorem of elementary calculus, a differentiable function f whose derivative is nonnegative on an interval I must be nondecreasing on that interval. This theorem has been generalized in many ways. For example, the differentiability of f has been replaced by a weaker regularity condition, the derivative has been replaced by various types of generalized derivative, and the set on which the derivative is assumed to exist, as well as the set on which it is assumed to be nonnegative, has been assumed to be less than all of the interval I. For example, the standard monotonicity theorem which appears in the theory of Lebesgue integration asserts that a function which is absolutely continuous and has a nonnegative derivative a.e. must be nondecreasing. A similar theorem involving the approximate derivative appears in connection with the integral of Denjoy-Khintchine.

In this section we consider several theorems whose conclusions are that a function is nondecreasing. We begin with a theorem of Goldowski [**54**] and Tonelli [**176**] (see also [**157**; p. 206]).

THEOREM. *Let f be a function satisfying the following conditions on the interval I:*

(i) *f is continuous,*
(ii) *f' exists (finite or infinite), except perhaps on a denumerable set,*
(iii) *$f' \geqq 0$ a.e.*

Then f is nondecreasing on I.

We note that condition (ii) cannot be weakened to the condition that the derivative exists except perhaps on a null set. This can be seen by considering the negative of the Cantor function.

In 1939 Tolstoff [**175**] obtained an improvement of the theorem of Goldowski-Tonelli.

THEOREM. *Let f be a function satisfying the following conditions on an interval I:*

(i) *f is approximately continuous,*
(ii) *f'_{ap} exists (finite or infinite) except perhaps on a denumerable set,*
(iii) *$f'_{ap} \geqq 0$ a.e.*

Then f is continuous and nondecreasing on I.

Another generalization of the Goldowski-Tonelli theorem was obtained by Zahorski [**192**: p. 19] in 1950.

THEOREM. *Let f be a function satisfying the following conditions on an interval I:*

(i) *f is a Darboux function,*
(ii) *f' exists (finite or infinite) except perhaps on a denumerable set,*
(iii) *$f' \geqq 0$ a.e.*

Then f is continuous and nondecreasing on I.

We note that Zahorski's Theorem is stronger than Tolstoff's in so far as Zahorski assumed only Darboux continuity instead of approximate continuity of f. On the other hand, his theorem is weaker in so far as conditions (ii) and (iii) involve the ordinary derivative instead of the approximate derivative. We would like a theorem which implies both Tolstoff's Theorem and Zahorski's Theorem. An obvious candidate for such a theorem is obtained by considering the weaker of the corresponding conditions of the two theorems. That is, must a Darboux function satisfying conditions (ii) and (iii) of Tolstoff's Theorem be nondecreasing? This question is answered in the negative, as can be seen by considering the example below. This example is a slight modification of an example found in [**191**: pp. 321, 322].

Example. Let f be a function satisfying the following conditions on the interval (0, 1).

(i) If (a, b) is an interval contiguous to the Cantor set then $f(a)=0$, $f(b)=1$ and f is continuous and nondecreasing on $[a, b]$.

(ii) If x is a two sided limit point of the Cantor set then $f(x)=1$.

(iii) Every two sided limit point of the Cantor set is a point of density of the set $\{x: f(x)=1\}$.

It is not difficult to verify that this function has the required properties.

So our first attempt to obtain a simultaneous generalization of the two theorems fails. What next? We note that hypothesis (i) of Tolstoff's Theorem implies that f be in Baire class 1. The same is true of hypothesis (ii) in Zahorski's Theorem. (The function in our example is in Baire class 2, but not in Baire class 1.) What happens if we add the requirement that f be in Baire class 1? That is, if f is a Darboux function in Baire class 1 and satisfies conditions (ii) and (iii) of Tolstoff's Theorem, must f be nondecreasing? This question was asked by Zahorski [**192**: p. 8]. It turns out that this question has an affirmative answer [**201**, **202**, **214**], thus providing a theorem which includes both the theorem of Tolstoff and the theorem of Zahorski. In fact, the following more general theorem is valid [**201**, **202**].

THEOREM. *Let $\mathcal{P}$ be a function-theoretic property sufficiently strong to imply*

(a) *Any Darboux function in Baire class* 1 *which satisfies property $\mathcal{P}$ on an interval I is VBG on I* [**157**: p. 221].

(b) *Any continuous function of bounded variation which satisfies property $\mathcal{P}$ on I is nondecreasing on I.*

Then any Darboux Baire 1 *function which satisfies property $\mathcal{P}$ on I is continuous and nondecreasing on I.*

To see that this theorem provides an affirmative answer to the question raised by Zahorski, we let $\mathcal{P}$ be the property of having, except perhaps on a denumerable set, an approximate derivative (finite or infinite) which is nonnegative a.e. Condition (a) follows from 10.8 of [**157**: p. 237] and condition (b) is a consequence of Tolstoff's Theorem.

Roughly speaking, the theorem states that if one wishes to show that a condition is strong enough to guarantee that every Darboux Baire 1 function satisfying the condition is nondecreasing, one need only show that every continuous function of bounded variation which satisfies the condition is nondecreasing. (Condition (a) is likely to be satisfied if the condition is at all "reasonable.") For example, one can use the theorem to show that Tolstoff's Theorem remains valid if approximate continuity is replaced by preponderant continuity and the approximate derivative is replaced by the preponderant derivative [**207**].

We conclude this section with two theorems concerning Dini derivatives.

Let f be a function defined on $[a, b]$ which satisfies
(a) $\limsup_{\xi\to x-} f(\xi) \leqq f(x) \leqq \limsup_{\xi\to x+} f(\xi)$ *for all* $x\in[a, b]$,
(b) $D^+f\geqq 0$ a.e. *on* $[a, b]$, *and*
(c) $D^+f> -\infty$ *except possibly on a denumerable set.*
Then f is nondecreasing.

This theorem is due to Gál [**41**]. Once more, none of the hypotheses of the theorem can be deleted with the conclusion still valid. Other theorems of this type, dealing with continuous functions, have been advanced by Garg [**43**] and Ważewski [**183**]. We state one such [**43**]: *Let f be a continuous function fulfilling Banach's condition* (T_2) (see Section 9). *Let* $Q=\{x: D^+f(x)<0\}$. *If* $mf(Q)=0$, *then f is nondecreasing.*

14. Derivatives and Darboux functions of Baire class 1. It was mentioned in Section 2 that every derivative belongs to Baire class 1 and possesses the Darboux property. The converse is not valid. Thus the requirement that a function f be a derivative is more stringent than the requirement that f be a Darboux Baire class 1 function. How much more stringent? On the one hand, Maximoff [**120**] has shown that the derivatives and the functions in Darboux Baire class 1 are topologically equivalent in the sense that any Darboux Baire class 1 function defined on $[a, b]$ can be transformed into a derivative by suitably transforming $[a, b]$ onto itself topologically. (See also [**24, 113**].) On the other hand, the two classes of functions exhibit quite different properties. We turn now to a consideration of some of these differences.

For α a real number and f any function defined on $[a, b]$ let $E_\alpha(f)=\{x: f(x)>\alpha\}$ and $E^\alpha(f)=\{x: f(x)<\alpha\}$. We mentioned in Section 4 that Zahorski considered such sets in trying to characterize derivatives. He was able to show that a necessary and sufficient condition that f be a Darboux Baire class 1 function is that each such set be an F_σ set with the property that each point of the set be a bilateral point of condensation of the set (he called this condition M_1). On the other hand, a *necessary* condition for a function f to be a derivative (possibly infinite) of a continuous function is that for every α, $E_\alpha(f)$ and $E^\alpha(f)$ be sets satisfying the condition which he called M_2. (A set E satisfies M_2 if E is an F_σ, and every one-sided neighborhood of each point in E intersects E is a set of positive measure.) This condition is not sufficient. Thus, each set of the type $E_\alpha(f)$ and $E^\alpha(f)$ must be considerably "more dense" near its members in order for f

to be a derivative (possibly infinite) than in order for f to be merely a Darboux Baire class 1 function. By requiring even more density of the sets $E_\alpha(f)$ and $E^\alpha(f)$, Zahorski managed to find necessary conditions for a function to be a *finite* derivative (condition M_3). The analogous necessary condition for f to be a *bounded* derivative is the still more stringent density condition that the sets $E_\alpha(f)$ and $E^\alpha(f)$ all be M_4 sets (see Section 4 for a definition of M_3 and M_4 sets). A *sufficient* density condition that the bounded function f be a derivative is that for every α, every point of $E_\alpha(f)(E^\alpha(f))$ be a point of (unit) density of $E_\alpha(f)$ (resp. $E^\alpha(f)$). This amounts to saying that the function is approximately continuous. The converse is, of course, not true; there exist bounded derivatives which are not approximately continuous. However, a partial converse, which characterizes approximately continuous functions, has been given by Lipiński [**94**]: *The function f is approximately continuous if and only if for every a and b the function $f_{ab}(x) = \max\{a, \min[b, f(x)]\}$ is a derivative.*

We mentioned in Section 4 that the question of *characterizing* the class of derivatives in terms of the set E_α and E^α has not yet been resolved. In this connection it should be mentioned that for the case of bounded derivatives no characterization solely in terms of the structure of the individual sets E^α and E_α is possible. This can be seen in the following way. Zahorski showed [**192**: pp. 45–47] that there are functions f, which are *not* bounded derivatives, but such that for all α the sets $E^\alpha(f)$ and $E_\alpha(f)$ are M_4 sets. Thus, if there were a condition of the type desired, it would have to be *more* stringent than the condition M_4. On the other hand, on page 35 one finds the result that for every M_4 set E there exists a bounded derivative f and a number α such that $E = E^\alpha(f)$. Thus the desired condition *cannot* be more stringent than M_4.

Some additional results concerning the sets $E^\alpha(f)$ and $E_\alpha(f)$ and M_k sets, $k = 2, 3, 4$, can be found in Lipiński [**91**, **93**, **95**]. The results of Zahorski concerning the sets $E_\alpha(f)$ and $E^\alpha(f)$ for Darboux Baire 1 functions have been extended to more general spaces by Mišik [**129**].

Another kind of comparison involving convergent interval functions was advanced by Neugebauer [**135**]. In this article the author gives characterizations of each of the two classes presently under consideration. This is done in such a way as to allow an interesting comparison between the two classes. Two conditions are stated. The first one is necessary and sufficient for a function to be a Darboux function of Baire class 1, whereas the two conditions together are necessary and sufficient for a function to be a derivative. Thus, it is precisely the second condition which shows how much more stringent a requirement it is for a function f to be a derivative than it is for f to be a Baire class 1 Darboux function. A precise formulation of the relevent theorems would require more space than is appropriate here, so we omit the details.

As we saw in Section 11, a necessary and sufficient condition that a set E be stationary for the class of derivatives is that $\sim E$ have inner measure zero [**117**], whereas a necessary and sufficient condition that E be stationary for the class of Darboux Baire class 1 functions is that $\sim E$ be totally imperfect [**16**], that

is, that $\sim E$ contain no nonempty perfect subset. For purposes of comparison, we mention that every totally imperfect set must have zero inner measure, but the converse statement is false. It is possible, however, for a totally imperfect set to have positive outer measure. In fact, the interval $[a, b]$ can be decomposed into two non-overlapping totally imperfect sets [**83**; p. 422]. It is clear that each of these sets must have outer measure equal to $b-a$. Similarly, the determining sets for the class of derivatives are those whose complements have zero inner measure, whereas the only determining set for the Darboux Baire class 1 functions is the interval $[a, b]$.

The remaining comparisons involve the algebraic and topological structures of the two classes. We begin by observing that the sum of two derivatives is again a derivative. The corresponding statement is not valid for the Darboux functions of Baire class 1. Thus let $f(x)=\sin(1/x)$, $f(0)=1$ and let $g(x)=-\sin(1/x)$, $g(0)=1$. Then $(f+g)(x)=0$, $(f+g)(0)=2$. The functions f and g are in the required class, but their sum is not.

On the other hand, if f is a Baire class 1 Darboux function, then so is f^2. This follows from the facts that a continuous function of a Baire function preserves the Baire class and a continuous function of a Darboux function is again Darboux. But the corresponding statement is not valid for derivatives. In fact if f is a square summable derivative, then f^2 is a derivative if and only if

$$\lim_{h\to 0}\frac{1}{h}\int_x^{x+h}|f(t)-f(x)|^2\,dt = 0 \quad \text{for all } x.$$

(See Iosifescu [**66**].) Some interesting related results may be found in Hruška [**205**], Iosifescu [**206**], Iosifescu and Marcus [**69**], Neugebauer [**137**], Sélivanoff [**149**], Wilkosz [**187**] and Wolff [**188**].

For bounded functions we can state a simpler result: *If f and f^2 are bounded on $[a, b]$, then both functions are derivatives if and only if f is approximately continuous* [**187**]. These results show that the square of a derivative (even a bounded derivative) need not be a derivative. We also see that even though a continuous function of a Darboux Baire class 1 function is still in that class, the corresponding result for derivatives is not valid. A sufficient condition for such a composition can be found in Choquet [**24**; p. 89]: *If g is a lower semicontinuous derivative and f is a continuous function of bounded variation on $(-\infty, \infty)$, then the function $f \circ g$ is a derivative.* It is true that this theorem puts considerable restrictions on both f and g. It would be of interest to know just how much these restrictions can be weakened. To give some slight indication of the difficulties that arise if we put no restrictions (other than that of being a derivative) on g, we state the following result found in [**24**; p. 89]: *If f is nondecreasing and continuous, and such that for every bounded derivative g the function $f \circ g$ is still a derivative, then f must be linear.*

A certain other comparison has, to the best of our knowledge, not yet been resolved. Let $\{f_n\}$ be a sequence of continuous functions converging pointwise to a limit function f. We know that if each f_n is continuous and the convergence

is *uniform*, then f is also continuous. Uniform convergence is, of course, not *necessary* for the limit function to be continuous. It has been known for a long time that a necessary and sufficient condition for the limit function to be continuous is that the convergence be quasi-uniform (see Hahn [**56**] for the relevant definition and theorem). One might ask for the corresponding types of convergence for the class of derivatives and for the class of Darboux Baire 1 functions: If $\{f_n\}$ is a sequence of derivatives (respectively Darboux Baire 1 functions) converging pointwise to a limit function f, then what additional restriction on the convergence is necessary and sufficient to guarantee that f also be a derivative (respectively Darboux Baire 1 function)? In this connection we mention that uniform convergence is sufficient in each case, but not necessary. The proof for derivatives is straightforward, and the proof for Darboux Baire 1 functions can be found in [**14**]. The relevant kind of convergence for functions in Baire class α (for fixed α) has been obtained by Gagaeff [**40**]. The results of Oeconomidis [**141, 142**] are relevant to this question for derivatives.

We conclude with a precise statement of Maximoff's deep theorem which we mentioned at the beginning of this section. (See also [**24**; p. 90].)

THEOREM. *Let f be a finite Darboux function of Baire class* 1 *on the interval* $[0, 1]$. *Then there exists a strictly increasing continuous function g such that* $g(0) = 0$, $g(1) = 1$ *and $f \circ g$ is a derivative.* (See [**120, 122**].)

Marcus [**113**] and Lipiński [**89**] consider some consequences of this theorem.

In [**121**], Maximoff showed that the word "derivative" can be replaced by the words "approximately continuous function" in the conclusion of the theorem.

In writing this article the authors benefited from discussions and correspondences with several mathematicians. Particular thanks are due to Professor John Olmsted for discussions when the project was in its initial stages, and to Professor Solomon Marcus for many valuable suggestions concerning relevant articles of which the authors were originally unaware.

The first author was supported in part by NSF Grant GP 1592.
The second author was an NSF Cooperative Graduate Fellow.

References

1. P. Alexandroff, L'intégration au sens de M. Denjoy considérée comme recherche des fonctions primitives, Rec. Math. Soc. Math. Moscou, 31 (1924) 465–476.

2. ———, Über die Äquivalenz des Perronschen und des Denjoyschen Integralbegriffes, Math. Z., 20 (1924) 213–222.

3. R. Baire, Sur les fonctions d'une variable réelle, Annali di Matematica, (3) 3 (1899).

4. S. Banach, Sur les fonctions dérivées des fonctions mesurables, Fund. Math., 3 (1922) 128–132.

5. ———, Über die Baire'sche Kategorie gewisser Funktionenmengen, Studia Math., 3 (1931) 174–179.

6. A. Besicovitch, Diskussion der stetigen Funktionen im Zusammenhang mit der Frage über ihre Differentierbarkeit, Bull. Acad. Sci. de Russie, 19 (1925) 527–540.

7. R. P. Boas, A primer of real functions, Carus Mathematical Monograph No. 13, MAA, 1960.

8. N. Boboc and S. Marcus, Sur la détermination d'une fonction par les valeurs prises sur un certain ensemble, Ann. Sci. École Norm. Sup. (3), 76 (1959) 151–159.

9. K. Bögel, Über die mehrdimensionale Differentiation, Jber. Deutsch. Math.-Verein, 65 (1962) 45–71.

10. E. Borel, Modèles arithmétiques et analytiques de l'irréversibilité apparente, C.R. Acad. Sci. Paris, 145 (1912) 1148–1151.

11. A. Bruckner, On derivatives with a dense set of zeros, Rev. Math. Pures Appl., 10 (1965) 149–153.

12. ———, Stationary sets for certain classes of derivates of Darboux functions, Mich. Math. J., 11 (1964) 305–309.

13. A. Bruckner and J. Ceder, Darboux continuity, Jber. Deutsch. Math.-Verein., 67 (1965) 93–117.

14. A. Bruckner, J. Ceder, and M. Weiss, On uniform limits of Darboux functions, Colloq. Math., (to appear).

15. A. Bruckner, and J. Leonard, On differentiable functions having an everywhere dense set of intervals of constancy, Canad. Math. Bull., 8 (1965) 73–76.

16. ———, Stationary sets and determining sets for certain classes of Darboux functions, Proc. Amer. Math. Soc., 16 (1965) 935–940.

17. A. Brudno, Continuity and differentiability, Rec. Math. (Mat. Sbornik) N.S., 13(55) (1943) 119–134. (Russian; English summary.)

18. P. Bullen, Construction of primitives of generalized derivatives with applications to trigonometric series, Canad. J. Math., 13 (1961) 48–58.

19. P. Butzer, Beziehungen zwischen den Riemannschen, Taylorschen und gewöhnlichen Ableitungen reelwertiger Funktionen, Math. Ann., 144 (1961) 275–298.

20. P. Butzer and W. Kozakiewicz, On the Riemann derivatives for integrable functions, Canad. J. Math., 6 (1954) 572–581.

21. A. Calderón and A. Zygmund, Local properties of solutions of elliptic partial differential equations, Studia Math., 20 (1961) 171–225.

22. ———, A note on local properties of solutions of elliptic differential equations, Proc. Nat. Acad. Sci. U.S.A., 46 (1960) 1385–1389.

23. Cellerier, Bull. Sci. Math., (2) 14 (1890) 152.

24. G. Choquet, Application des propriétés descriptives de la fonction contingent à la théorie des fonctions de variable réelle et à la géométrie différentielle des variétés cartésiennes, J. Math. Pures Appl., (9) 26 (1947) 115–226.

25. S. Chow, On approximate derivatives, Bull. Amer. Math. Soc., 54 (1948) 793–802.

26. J. Clarkson, A property of derivatives, Bull. Amer. Math. Soc., 53 (1947) 124–125.

27. A. Császár, Sur une généralisation de la notion de dérivée, Acta Sci. Math., (Szeged) 16 (1955) 137–159.

28. A. Denjoy, Calcul de la primitive de la fonction dérivée la plus générale, C. R. Acad. Sci. Paris, 154 (1912) 1075–1078.

29. ———, Mémoire sur la totalisation des nombres dérivés non-sommables, Ann. École. Norm. Sup., 33 (1916) 127–236; 34 (1917) 181–238.

30. ———, Mémoire sur les nombres dérivés des fonctions continues, J. Math. Pures Appl., (7) 1 (1915) 105–240.

31. ———, Sur les fonctions dérivées sommables, Bull. Soc. Math. France, 43 (1915) 161–248.

32. ———, Sur l'intégration des coefficients différentiels d'ordre supérieur, Fund. Math., 25 (1935) 273–326.

33. ———, Sur une propriété des fonctions dérivées, Enseignement Math., 18 (1916) 320–328.

34. A. Denjoy, Une extension de l'intégrale de M. Lebesgue, C. R. Acad. Sci. Paris, 154 (1912) 859–862.

35. S. Eilenberg and S. Saks, Sur la dérivation des fonctions dans des ensembles dénombrables, Fund. Math., 25 (1935) 264–266.

36. M. Esser and O. Shisha, A modified differentiation, this MONTHLY, 71 (1964) 904–906.

37. H. Fast, Une remarque sur la propriété de Weierstrass, Colloq. Math., 7 (1959) 75–77.

38. F. Filipczak, On the derivative of a discontinuous function, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys., 12 (1964) 535–537.

39. M. Fréchet, Sur diverses définitions de la différentiabilité, Enseignement Math., (2) 10 (1964) 177–228.

40. B. Gagaeff, Sur les suites convergentes de fonctions mesurables B, Fund. Math., 18 (1932) 182–188.

41. I. Gál, On the fundamental theorems of the calculus, Trans. Amer. Math. Soc., 86 (1957) 309–320.

42. K. Garg, An analogue of Denjoy's theorem, Ganita, 12 (1961) 9–14.

43. ———, Applications of Denjoy analogue I. (Sufficient conditions for a function to be monotone), Ann. Polon. Math., 15 (1964) 159–165.

44. ———, Applications of Denjoy analogue II. (Local structure of level sets and Dini derivates), Acta Math. Acad. Sci. Hungar., 14 (1963) 183–186.

45. ———, Applications of Denjoy analogue III. (Distribution of various typical level sets), Acta Math. Acad. Sci. Hungar., 14 (1963) 187–195.

46. ———, On nowhere monotone functions, I (Derivates at a residual set), Ann. Univ. Sci. Budapest, Eötvös Sect. Math., 5 (1962) 173–177.

47. ———, On nowhere monotone functions, II (Derivates at sets of power c and at sets of positive measure), Rev. Math. Pures Appl., 7 (1962) 663–671.

48. ———, On nowhere monotone functions, III (Functions of the first and second species), Rev. Math. Pures Appl., 8 (1963) 83–90.

49. ———, On the derivability of functions discontinuous at a dense set, Rev. Math. Pures Appl., 7 (1962) 175–179.

50. B. Gelbaum and J. Olmsted, Counterexamples in analysis, Holden-Day, San Francisco, 1964.

51. C. Goffman, Real functions, Rinehart, New York, 1960.

52. C. Goffman and C. Neugebauer, Linear continuous functions, Proc. Amer. Math. Soc., 12 (1961) 997–998.

53. ———, On approximate derivatives, Proc. Amer. Math. Soc., 11 (1960) 962–966.

54. G. Goldowski, Note sur les dérivées exactes, Rec. Math. Soc. Math. Moscou, (Mat. Sbornik), 35 (1928) 35–36.

55. E. Görlich and R. Nessel, Über Peano-Ableitungen in der Norm (to appear).

56. H. Hahn, Reele Funktionen, Chelsea, New York, 1958.

57. ———, Über den Fundamentalsatz der Integralrechnung, Monatshefte Math. Phys., 16 (1905) 161–166.

58. O. Hájek, Note sur la mesurabilité B de la dérivée supérieure, Fund. Math., 44 (1957) 238–240.

59. H. Hake, Über de la Vallée Poussins Ober- und Unterfunktionen einfacher Integral und die Integraldefinition von Perron, Math. Ann., 83 (1921) 119–142.

60. I. Halperin, Discontinuous functions with the Darboux property, this MONTHLY, 57 (1950) 539–540.

61. ———, Discontinuous functions with the Darboux property, Canad. Math. Bull., 2 (1959) 111–118.

62. F. Hausdorff, Set theory, Chelsea, New York, 1957.

63. E. Hobson, Theory of functions of a real variable, Vol. 1, Dover, New York, 1957.

64. ———, Theory of functions of a real variable, Vol. 2, Dover, New York, 1957.

65. F. Hsiang, On differentiable functions, Bull. Amer. Math. Soc., 66 (1960) 382–383.

66. M. Iosifescu, Conditions that the product of two derivatives be a derivative, Rev. Math. Pures Appl., 4 (1959) 641–649. (In Russian.)

67. ———, Sur les fonctions continues dont les ensembles de niveau sont au plus dénombrables, Rev. Math. Pures Appl., 3 (1958) 439–441.

68. ———, Sur un théorème de A. Marchaud, Com. Acad. R. P. Romîne, 6 (1956) 1169–1171. (Romanian; Russian and French summaries.)

69. M. Iosifescu and S. Marcus, Sur un problème de P. Scherk, concernant la somme des carrés de deux dérivées, Canad. Math. Bull., 5 (1962) 129–132.

70. V. Jarnik, Tohoku Math. J., 37 (1933) 248.

71. R. Jeffery, The theory of functions of a real variable, Mathematical Expositions No. 6, University of Toronto Press, 1962.

72. C. Kassimatis, Functions which have generalized Riemann derivatives, Canad. J. Math., 10 (1958) 413–420.

73. A. Khintchine, Recherches sur la structure des fonctions mesurables, Fund. Math., 9 (1927) 212–279.

74. ———, Recherches sur la structure des fonctions mesurables, Recueil Mathém., 31 (1924) 265–285, 377–433. (Russian.)

75. ———, Sur le procédé d'intégration de M. Denjoy, Rec. Math. Soc. Math. Moscou, 30 (1918) 543–557.

76. ———, Sur une extension de l'intégrale de M. Denjoy, C. R. Acad. Sci. Paris, 162 (1916) 287–291.

77. A. Köpcke, Über Differentierbarkeit und Anschaulichkeit der stetigen Funktionen, Math. Ann., 29 (1887) 123–140.

78. ———, Über eine durchaus differentierbare stetige Funktion mit Oscillationen in jedem Intervalle, Math. Ann., 34 (1889) 161–171.

79. ———, Nachtrag zu dem Aufsatze "Über eine durchaus differentierbare stetige Funktion mit Oscillationen in jedem Intervalle," Math. Ann., 35 (1890) 104–109.

80. A. Kronrod, Sur la structure de l'ensemble des points de discontinuité d'une fonction dérivable en ses points de continuité, Bull. Acad. Sci. URSS Sér. Math., (1939) 569–578.

81. K. Krzyzewski, Note on approximate derivatives, Colloq. Math., 10 (1963) 281–285.

82. M. Kulbacka, Sur certaines propriétés des dérivées approximatives, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys., 12 (1964) 17–20.

83. C. Kuratowski, Topologie I, 4th ed., Monografie Matematyczne 20, Warszawa, 1958.

84. ———, Topologie II, 3rd ed., Monografie Matematyczne 21, Warszawa, 1961.

85. ——— and W. Sierpiński, Les fonctions de classe 1 et les ensembles connexes punctiformes, Fund. Math., 3 (1922) 303–313.

86. B. Kuttner, Some theorems on fractional derivatives, Proc. London Math Soc., (3) 3 (1953) 480–497.

87. E. Landis, On the set of points of existence of an infinite derivative, Dokl. Akad. Nauk SSSR (N.S.), 107 (1956) 202–204. (Russian.)

88. J. Liberman, Théorème de Denjoy sur la dérivée d'une fonction arbitraire par rapport à une fonction continue, Rec. Math. (Mat. Sbornik) N. S., 9 (1941) 221–236.

89. J. Lipiński, Mesure et dérivée, Colloq. Math., 8 (1961) 83–88.

90. ———, Sur les dérivées de Pompeiu, Rev. Math. Pures Appl., 10 (1965) 447–451.

91. ———, Sur certains problèmes de Choquet et de Zahorski concernant les fonction dérivées, Fund. Math., 44 (1957) 94–102.

92. ———, Sur la discontinuité approximative et la dérivée approximative, Colloq. Math., 10 (1963) 103–109.

93. ———, Sur les ensembles $\{f'(x)>a\}$, Fund. Math., 45 (1958) 254–260.

94. ———, Sur les fonctions approximativement continues, Colloq. Math., 5 (1958) 172–175.

95. ———, Une propriété des ensembles $\{f'(x)>a\}$, Fund. Math., 42 (1955) 339–342.

96. H. Looman, Über die Perronsche Integraldefinition, Math. Ann., 93 (1925) 153–156.

97. N. Lusin, Sur la notion de l'intégrale, Annali Mat. Pura et Appl., (3) 26 (1917) 77–129.

98. ———, Sur les propriétés de l'intégrale de M. Denjoy, C. R. Acad. Sci. Paris, 155 (1912) 1475–1478.

99. ———, Trigonometric integrals and series, Moscow, 1915. (Russian.)

100. A. Marchaud, Sur une condition de quasi-rectificabilité, Fund. Math., 20 (1933) 105–116.

101. J. Marcinkiewicz, Sur les nombres dérivés, Fund. Math., 24 (1935) 305–308.

102. ———, Sur quelques intégrales du type de Dini, Ann. de la Soc. Pol. de Math., 17 (1938) 42–50.

103. J. Marcinkiewicz and A. Zygmund, On the differentiability of functions and summability of trigonometrical series, Fund. Math., 26 (1936) 1–43.

104. S. Marcus, Les ensembles stationnaires de certaines classes de fonctions, C. R. Acad. Sci. Paris, 254 (1962) 1186–1188.

105. ———, Les ensembles stationnaires de certaines classes de fonctions dérivées, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., 32 (1962) 484–487.

106. ———, Les fonctions de Pompeiu, Acad. R. P. Romîne Stud. Cerc. Mat., 5 (1954) 413–419. (Romanian; Russian and French summaries.)

107. ———, Functions with the Darboux property and functions with connected graphs, Math. Ann., 141 (1960) 311–317.

108. ———, On a theorem of Denjoy and on approximate derivatives, Monatsh. Math. 66 (1962) 435–440.

109. ———, Points of discontinuity and points at which the derivative is infinite, Rev. Math. Pures Appl., 7 (1962) 309–318. (Russian.)

110. ———, Remarques sur les fonctions intégrables au sens de Riemann, Bull. Math. Soc. Sci. Math. Phys. R. P. Roumaine, 2 (50) (1958) 433–439.

111. ———, Sur la dérivée approximative qualitative, Com. Acad. R. P. Romîne, 3 (1953) 361–364.

112. ———, Sur la limite approximative qualitative, Com. Acad. R. P. Romîne, 3 (1953) 9–12.

113. ———, Sur les dérivées dont les zéros forment un ensemble frontière partout dense, Rend. Circ. Mat. Palermo, (2) 12 (1963) 1–36.

114. ———, Sur les ensembles déterminant des dérivées approximatives, C. R. Acad. Sci. Paris, 255 (1962) 1685–1687.

115. ———, Sur les ensembles stationnaires de fonctions dérivées—finies ou infinies, Com. Acad. R. P. Romîne, 12 (1962) 399–402. (Romanian; Russian and French summaries.)

116. ———, Sur un problème de Z. Zahorski concernant les points où la dérivée est infinie, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., 29 (1960) 176–180.

117. ———, Sur une généralisation de la notion de quasi-analyticité, C. R. Acad. Sci. Paris, 254 (1962) 985–987.

118. E. Marczewski, Remarks on sets of measure zero and the derivability of monotonic functions, Prace Mat., 1 (1955) 141–144. (Polish; English and Russian summaries.)

119. A. Matysiak, O granicach i pochodnych aproksymatywnch, thesis, Łódz, 1960. (Polish.)

120. I. Maximoff, On continuous transformation of some functions into an ordinary derivative, Ann. Scuola Norm. Sup. Pisa., 12 (1943) 147–160.

121. ———, Sur la transformation continue de fonctions, Bull. Soc. Phys. Math. Kazan, (3) 12 (1940) 9–41. (Russian; French summary.)

122. ———, Sur la transformation continue de quelques fonctions en dérivées exactes, Bull. Soc. Phys Math. Kazan, (3) 12 (1940) 57–81. (Russian; French summary.)

123. S. Mazurkiewicz, Sur les fonctions non-dérivables, Studia Math., 3 (1931) 92–94.

124. ———, Sur les nombres dérivés, Fund. Math., 23 (1934) 9–10.

125. K. Menger, Multiderivatives and multi-integrals, this MONTHLY, 65 (1957) 58–70.

126. ———, Rates of change and derivatives, Fund. Math., 46 (1958) 89–102.

127. K. Menger and S. Shu, Generalized derivatives and expansions, Proc. Nat. Acad. Sci. U.S.A., 41 (1955) 591–595.

128. S. Minetti, Sull'operazione di derivazione, Atti Acad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., (8) 8 (1950) 27–31.

129. L. Mišik, Über die Funktionen der ersten Baireschen Klasse mit der Eigenschaft von Darboux, Mat. Fyz. Časopis Sloven. Akad. Vied., 14 (1964) 44–49.

130. A. Morse, Dini derivatives of continuous functions, Proc. Amer. Math. Soc., 5 (1954) 126–130.

131. P. Murav'ev, A generalized derivative and its application to ordinary differential equations, Izv. Vysš. Učebn. Zaved. Matematika, 1 (26) (1962) 89–100. (Russian.)

132. I. Natanson, Theory of functions of a real variable, Vol. 1, Ungaı, New York, 1961.

133. ———, Theory of functions of a real variable, Vol. 2, Ungar, New York, undated.

134. C. Neugebauer, A class of functions determined by dense sets, Arch. Math., 12 (1961) 206–209.

135. ———, Darboux functions of Baire class 1 and derivatives, Proc. Amer. Math. Soc., 13 (1962) 838–843.

136. ———, Differentiability almost everywhere (to appear).

137. ———, On a paper by M. Iosifescu and S. Marcus, Canad. Math. Bull., 6(1963) 367–371.

138. ———, Smoothness and differentiability in L_p, Studia Math., 25 (1964) 81–91.

139. ———, Symmetric, continuous, and smooth functions, Duke Math. J., 31 (1964) 23–32.

140. ———, A theorem on derivatives, Acta Sci. Math. (Szeged), 23 (1962) 79–81.

141. N. Oeconomidis, Sur les nombres dérivés d'une suite de fonctions réelles, C. R. Acad. Sci. Paris, 256 (1963) 3229–3232.

142. ———, Sur les valeurs limites et les nombres dérivés d'une suite de fonctions réelles, C. R. Acad. Sci. Paris, 256 (1963) 1208–1211.

143. H. Oliver, The exact Peano derivative, Trans. Amer. Math. Soc., 76 (1954) 444–456.

144. A. O'Neill, Contributions to the theory of derivatives, Duke Math. J., 12 (1945) 89–99.

145. G. Peano, Sulla formula di Taylor, Atti della Accad. delle Scienze di Torino, 27 (1891) 40–46. (=*Opere scelte*, V. 1, Edizioni Cremonese, Roma, 1957, pp. 204–209.)

146. ———, Sur la définition de la dérivée, Mathesis, (2) 2 (1892) 12–14. (=*Opere scelte*, V. 1, Edizioni Cremonese, Rome, 1957, pp. 210–212.)

147. E. Pepper, On continuous functions without a derivative, Fund. Math., 12 (1928) 244–263.

148. O. Perron, Über den Integralbegriff, S. B. Heidelberg, Akad. Wiss., 16 (1914).

149. B. Pettineo, Quelques observations sur les fonctions dérivables presque partout, C. R. Acad. Sci. Paris, 248 (1959) 518–520.

150. ———, Sulla derivazione delle funzioni continue, Atti Accad. Sci. Lett. Arti Palermo, Parte I, 4 (1956/7) 211–238.

151. D. Pompeiu, Sur les fonctions dérivées, Math. Ann., 63 (1906) 326–332.

152. K. Popoff, Über die verallgemeinerten Ableitungen, die durch ein Iterationsverfahren gebildet sind, Abh. Preuss. Akad. Wiss. Math. Nat. Kl., 1942, No. 2, 19 pp.

153. R. Rinehart and J. Wilson, Two types of differentiability of functions on algebras, Rend. Circ. Mat. Palermo, (2) 11 (1962) 204–216.

154. S. Ruziewicz, Sur les fonctions qui ont la même dérivée et dont la différence n'est pas constante, Fund. Math., 1 (1920) 148–151.

155. S. Saks, On the functions of Besicovitch in the space of continuous functions, Fund. Math., 19 (1932) 211–219.

156. ———, Sur les nombres dérivés des fonctions, Fund. Math., 5 (1924) 98–104.

157. ———, Theory of the integral, Monografie Matematyczne 7, Warszawa-Lwów, 1937.

158. W. Sargent, The Borel derivatives of a function, Proc. London Math. Soc., (2) 38 (1935) 180–196.

159. N. Sélivanoff, Note sur les fonctions dérivées, Bull. Math. Mech. Inst. Univ. Tomsk, 3 (1946) 125–127.

160. U. Shukla, On points of non-symmetrical differentiability of a continuous function I., Ganita, 2 (1951) 54–61.

161. U. Shukla, On points of non-symmetrical differentiability of a continuous function II., Ganita, 4 (1953) 139–141.

162. W. Sierpiński, Sur les fonctions dérivées des fonctions discontinues, Fund. Math., 3 (1922) 123–127.

163. ———, Sur une propriété de fonctions quelconques d'une variable réelle, Fund. Math., 25 (1935) 1–4.

164. G. H. Sindalovskiĭ, Congruent and asymptotic differentiability, Dokl. Akad. Nauk, SSSR, 150 (1963) 995–997. (Russian.)

165. ———, Continuity and differentiability with respect to congruent sets, Dokl. Akad. Nauk SSSR, 134 (1960) 1305–1306. (Russian; translated in Soviet Math. Dokl., 1 (1961) 1217–1218.)

166. ———, On a generalization of derived numbers, Izv. Akad. Nauk SSSR Ser. Mat., 24 (1960) 707–720. (Russian.)

167. A. Singh, On infinite derivatives, Fund. Math., 33 (1945) 106–107.

168. J. Staniszewska, Sur la classe de Baire des dérivées de Dini, Fund. Math., 47 (1959) 215–217.

169. E. Stein and A. Zygmund, On the differentiability of functions, Studia Math., 23 (1964) 247–283.

170. ———, Smoothness and differentiability of functions, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., 3–4 (1960–61) 295–307.

171. F. Sunyer Balaguer, Sobre la determinación de una función mediante sus números derivados, Collect. Math., 10 (1958) 185–194.

172. ———, Sur la détermination d'une fonction par ses nombres dérivés, C. R. Acad. Sci. Paris, 245 (1957) 1690–1692.

173. H. Thielman, Theory of functions of a real variable, Prentice-Hall, Englewood Cliffs, N. J., 1953.

174. G. Tolstoff, Sur la dérivée approximative exacte, Rec. Math. (Mat. Sbornik) N.S., 4 (1938) 499–504.

175. ———, Sur quelques propriétés des fonctions approximativement continues, Rec. Math. (Mat. Sbornik) N.S., 5 (1939) 637–645.

176. L. Tonelli, Sulle derivate esatte, Memorie della Accademia delle Scienze dell'Instituto Bologna, (8) 8 (1930–31) 13–15.

177. V. Tzodiks, On sets of points where the derivative is $+\infty$ or $-\infty$ correspondingly, Dokl. Akad. Nauk, SSSR (N.S.), 113 (1957) 36–38. (Russian.)

178. ———, On sets of points where the derivative is finite or infinite correspondingly, Dokl. Akad. Nauk, SSSR (N.S.), 114 (1957) 1174–1176. (Russian.)

179. ———, On sets of points where the derivative is equal to $+\infty$ or $-\infty$ respectively, Mat. Sbornik. N.S., 43 (85) (1957) 429–450. (Russian.)

180. C. de la Vallée Poussin, Intégrales de Lebesgue. Fonctions d'ensembles. Classes de Baire, Gauthier-Villars, Paris, 1916.

181. ———, Sur l'approximation des fonctions d'une variable réelle et leurs dérivées par les polynômes et les suites limitées de Fourier, Bull. Acad. Royale de Belgique, (1908), 193–254.

182. V. Volterra, Sui principii del calcolo integrale, Giorn. di Battaglini, 19 (1881) 333–372.

183. T. Ważewiski, Sur une condition nécessaire et suffisante pour qu'une fonction continue soit monotone, Ann. Soc. Polonaise Math., 24 (1951) 111–119.

184. K. Weierstrass, Zur Funktionenlehre, Monatsb. Akad. Wiss. Berlin, August 1880, 719–743. (In Werke, V. 2.)

185. C. Weil, On properties of derivatives, Trans. Amer. Math. Soc., 114 (1965) 363–376.

186. M. Weiss, On symmetric derivatives in L^p, Studia Math., 24 (1964) 89–100.

187. W. Wilkosz, Some properties of derivative functions, Fund. Math., 2 (1921) 145–154.

188. J. Wolff, On derived functions of a real variable, Proc. Royal Acad. Amsterdam, 28 (1924) 282–285.

189. G. Young, On the derivates of a function, Proc. London Math. Soc., (2) 15 (1916) 360–384.

190. Z. Zahorski, Punktmengen in welchem eine stetige Funktion nicht differenzierbar ist, Rec. Math. (Mat. Sbornik), (9) 51 (1941) 487–510. (Russian; German summary.)

191. ———, Sur la classe de Baire des dérivées approximatives d'une fonction quelconque, Ann. Soc. Polon. Math., 21 (1948) 306–323.

192. ———, Sur la première dérivée, Trans. Amer. Math. Soc., 69 (1950) 1–54.

193. ———, Sur l'ensemble des points de non-dérivabilité d'une fonction continue, Bull. Soc. Math. France, 74 (1946) 147–178.

194. ———, Über die Konstruktion einer differenzierbaren monotonen, nicht konstanten Funktion mit überall dichter Menge von Konstanzintervallen, C. R. Société de Sciences et Lettres de Varsovie, Classe III, 30 (1937) 202–206.

195. ———, Über die Menge der Punkte in welchen die Ableitung unendlich ist, Tohoku Math. J., 48 (1941) 321–330.

196. Z. Zalcwasser, Sur les fonctions de Köpcke, Prace Mat. Fiz., 35 (1927–28) 57–99. (Polish; French summary.)

197. A. Zygmund, Smooth functions, Duke Math. J., 12 (1945) 47–76.

198. ———, Trigonometric Series, Vol. 1, University Press, Cambridge, 1959.

199. ———, Trigonometric Series, Vol. 2, University Press, Cambridge, 1959.

SUPPLEMENTARY REFERENCES

200. B. Bojarski, Sur la dérivée d'une fonction discontinue, Ann. Soc. Polon. Math., 24 (1953) 190–191.

201. A. Bruckner, An affirmative answer to a problem of Zahorski and some consequences, Mich. Math. J. (to appear).

202. ———, A theorem on monotonicity and a solution to a problem of Zahorski, Bull. Amer. Math. Soc., 71 (1965) 713–716.

203. ——— and C. Goffman, The boundary behavior of real functions in the upper half plane, Rev. Math. Pures Appl. (to appear).

204. F. Filipczak, On the derivative of a discontinuous function, Colloq. Math., 13 (1964) 73–79.

205. V. Hruška, Une note sur les fonctions aux valeurs intermédiaires, Časopis Pěst. Mat., 71 (1946) 67–69.

206. M. Iosifescu, On the product of two derivatives, Com. Acad. R. P. Romîne, 7 (1957) 319–321. (Romanian.)

207. J. Leonard, Some conditions implying the monotonicity of a real function (to appear).

208. J. Lipiński, Sur quelques problèmes de S. Marcus relatifs à la dérivée d'une fonction monotone, Rev. Math. Pures Appl., 8 (1963) 449–454.

209. M. Neubauer, Über die partiellen Derivierten unstetiger Funktionen, Monatsh. Math., 38 (1931) 139–146.

210. G. Piranian, The derivative of a monotonic discontinuous function, Proc. Amer. Math. Soc., 16 (1965) 243–244.

211. ———, The set of nondifferentiability of a continuous function, (to appear).

212. G. Sindalovskiĭ, Differentiability with respect to congruent sets, Izv. Akad. Nauk SSSR Ser. Math., 29 (1965) 11–40.

213. W. Stepanoff, Sur les conditions de l'existence de la différentielle totale, Rec. Math. Soc. Math. Moscou (Mat. Sb.), 32 (1925) 511–526.

214. T. Świątkowski, On the conditions of monotonicity of functions, Fund. Math., 59 (to appear).

215. M. Weiss, Total and partial differentiability in L_p, Studia Math., 25 (1964) 103–109.

216. H. Whitney, On totally differentiable and smooth functions, Pacific J. Math., 1 (1951) 143–159.

THE SET OF NONDIFFERENTIABILITY OF A CONTINUOUS FUNCTION

GEORGE PIRANIAN, University of Michigan

In this note, I give a simple proof of the following proposition, which (together with its converse) was first established by Z. Zahorski [**3**, pp. 492–507].

THEOREM (Zahorski). *If E is the union of two sets on the real line, the first of type G_δ, the second of type $G_{\delta\sigma}$ and of measure 0, then there exists a continuous function whose derivative exists nowhere on E but everywhere on the complement of E.*

(A set is of type G_δ provided it is the intersection of countably many open sets; without loss of generality, we may assume that the open sets form a decreasing sequence. A set is of type $G_{\delta\sigma}$ if it is the union of countably many sets of type G_δ.)

Zahorski's theorem is an obvious consequence of the two lemmas below. We formulate the lemmas in terms of the Dini derivatives D^+f, D_+f, D^-f, D_-f. Of these, the upper and lower right-hand derivatives, for example, are defined as

$$D^+f(x) = \limsup_{h\to+0}\frac{f(x+h)-f(x)}{h}, \qquad D_+f(x) = \liminf_{h\to+0}\frac{f(x+h)-f(x)}{h}.$$

LEMMA 1. *If E_0 is a set of type G_δ on the real line, then there exists a uniformly continuous function, differentiable everywhere outside of E_0, whose set of Dini derivatives includes the values ∞ and $-\infty$, at each point of E_0.*

LEMMA 2. *If E is a set of type G_δ and measure 0 on the real line, then there exists a continuous function f such that*

(i) $|f(x+h)-f(x)| \leqq |h|$ *for all x and h,*

(ii) $f'(x)$ *exists whenever $x \notin E$,*

(iii) *if $x\in E$, then either $D^+f(x)-D_+f(x)\geqq K$ or $D^-f(x)-D_-f(x)\geqq K$, where K is a positive constant independent of E.*

To see that the lemmas imply the theorem, suppose that $E=\bigcup_0^\infty E_j$, where each set E_j is of type G_δ and (for $j\geqq 1$) of measure 0. Suppose further that the function f_0 has the properties described in Lemma 1, and that for $j=1, 2, \cdots$ the function f_j has the properties described in Lemma 2, relative to E_j. If α is a sufficiently small positive constant, then the function $f(x)=\sum \alpha^j f_j(x)$ is continuous and has the set E as its set of nondifferentiability.

This proof does not differ basically from Zahorski's proof. Our Lemma 1 is essentially equivalent to Zahorski's Theorems I and II [**3**, p. 493]. Our Lemma 2 is equivalent to Lemma III on page 504 of [**3**] (Zahorski establishes uniform nondifferentiability on E, although he does not claim it in the statement of his lemma). The justification of the present note therefore lies not in a new approach to the theorem, but in simpler proofs of the auxiliary propositions. Instead of carrying out numerical computations, we describe our constructions in such a way that a diligent student can supply the technical details.

The motivation for our proof of Lemma 1 lies with the continuous but nowhere differentiable function constructed by Weierstrass [1, pp. 29–31], [2, pp. 97–100].

By hypothesis, E_0 is the intersection of a decreasing sequence $\{G_n\}_1^\infty$ of open sets. For the sake of simplicity, we suppose first that E_0 contains no interval of infinite length. Without loss of generality, we may then assume that each component of G_n is a finite interval (a_{nk}, b_{nk}). In the complement of G_n, let $g_n(x) = 0$. In (a_{nk}, b_{nk}), let

$$g_n(x) = c_{nk}(x - a_{nk})^2(x - b_{nk})^2 \cos \frac{d_{nk}}{(x - a_{nk})(x - b_{nk})},$$

where $c_{nk} = 4^{-n} \min [1, (b_{nk} - a_{nk})^{-4}]$, and where d_{nk} denotes a positive constant. Since

$$\frac{d}{dx} \frac{1}{(x - a)(x - b)} = \frac{a + b - 2x}{(x - a)^2(x - b)^2},$$

the shape of the waves in the graph of g_n approaches a limiting shape, near the endpoints of the interval (a_{nk}, b_{nk}). The ratio between the height and the length of each wave is approximately $(b_{nk} - a_{nk})c_{nk}d_{nk}/2\pi$.

Now let $f(x) = \sum g_n(x)$. By the choice of c_{nk}, the waves of $g_{n+1}, g_{n+2}, \cdots$ can not mask the waves of g_n. It follows that if we choose the d_{nk} large enough, then every neighborhood of each point x_0 in $\bigcap G_n$ contains points x_1 and x_2 ($x_1 < x_0 < x_2$) such that the difference quotients

$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} \quad \text{and} \quad \frac{f(x_2) - f(x_0)}{x_2 - x_0}$$

are arbitrarily large and have opposite signs. Consequently, f is not differentiable at any point of E. On the other hand, each function g_n is differentiable everywhere, and an elementary argument shows that f is differentiable everywhere outside of E.

If E contains an infinite interval, suppose (for example) that the maximal infinite interval of E toward the right is (a, ∞) or $[a, \infty)$. For $a < x < \infty$, we modify the previous definition: we now write

$$f(x) = \begin{cases} (x - a)^2 \sum 4^{-n} \cos 4^{2n}/(x - a) & \text{in case } a \notin E, \\ \sqrt{x - a} \sum 4^{-n} \cos 4^{2n}/(x - a) & \text{in case } a \in E. \end{cases}$$

Then f has the required differentiability property. To achieve uniform continuity of f, we further replace the quantity $x - a$ in the two alternate expressions with $(x - a)/[1 + (x - a)^2]$.

To prove Lemma 2, let $E = \bigcap G_n$ $(G_n \supset G_{n+1})$, denote by $\{(a_{nk}, b_{nk})\}$ the family of components of G_n, and define

$$h_1(x) = \begin{cases} \dfrac{(x - a_{1k})^2(x - b_{1k})^2}{(b_{1k} - a_{1k})^3} & (a_{1k} < x < b_{1k}), \\ 0 & (x \notin E_1). \end{cases}$$

Over each component of E_1, the graph of h_1 shows a mound. All mounds have the same shape (see Figure 1; the vertical scale is exaggerated). At the second

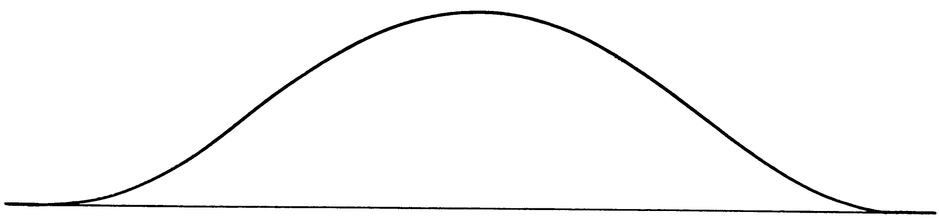


FIG. 1

stage, we shall create a similar mound over each component of E_2. However, in order to ensure boundedness of the difference quotients, we first modify the mounds of the first stage so that each mound of the second stage rests again on a horizontal segment. To this purpose, we cover each component (a_{2k}, b_{2k}) of G_2 symmetrically with a segment I_{2k} of length $2(b_{2k}-a_{2k})$. (The segments I_{2k} are not necessarily disjoint.) For each k, let ϕ_{2k} be continuous, vanish in (a_{2k}, b_{2k}), take the value 1 outside of I_{2k}, and be linear in the two segments of I_{2k} that protrude beyond (a_{2k}, b_{2k}). We define

$$\phi_2(x) = \inf_k \phi_{2k}(x), \qquad h_1^*(x) = \int_0^x \phi_2(t) h_1'(t)\, dt.$$

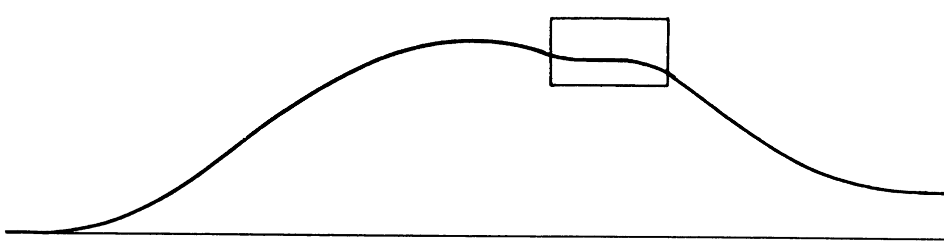


FIG. 2

The function h_1^* is differentiable, and its graph differs from the graph of h_1 only in being flat over G_2 (with all corners rounded off) so that the mounds are slightly distorted (see Figure 2). The distortion may shift the two endpoints of some mounds to different heights; consequently, the function h_1^* is not necessarily constant in the complement of E_1. However, since E has measure 0, we

may suppose the measure of E_2 to be as small as we like, in each component of E_1, so that the distortions are correspondingly small.

On each component (a_{2k}, b_{2k}) of G_2, we define η_2 by the formula

$$\eta_2(x) = \frac{(x - a_{2k})^2(x - b_{2k})^2}{(b_{2k} - a_{2k})^3},$$

and outside of E_2 we set $\eta_2(x) = 0$. Let $h_2(x) = h_1^*(x) + \eta_2(x)$ (see Figure 3), level the graph of $h_2(x)$ in the set G_3, denote the new function by h_2^*, raise new mounds over G_3, and continue indefinitely.

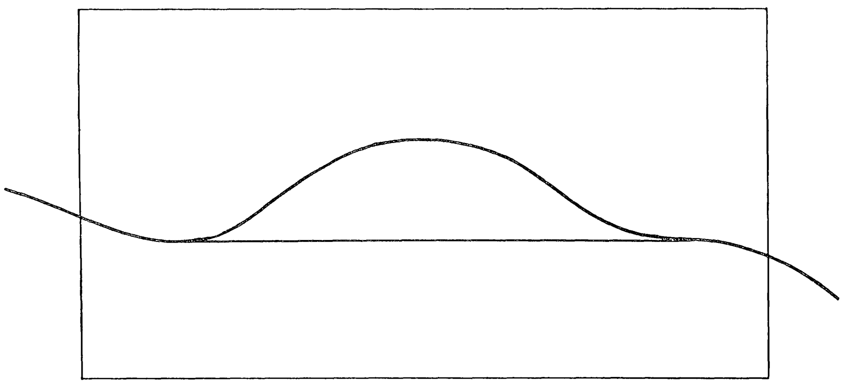


FIG. 3

Let $f(x) = \lim h_n(x)$. If $x_0 \in E$, then the point $(x_0, f(x_0))$ lies on a mound of the nth stage, for $n = 1, 2, \cdots$. For each n, the corresponding mound has two points to the left or two points to the right of x_0 such that the corresponding difference quotients differ by at least a certain positive universal constant K.

Clearly, f satisfies conditions (i) and (iii) in Lemma 2. It does not necessarily satisfy condition (ii). Suppose, for example, that x_0 does not belong to G_1 but is a limit point of a sequence of components of G_1. In that case, it may happen that among the line segments joining the point $(x_0, 0)$ to the highest points on the corresponding mounds, infinitely many have slopes of absolute value greater than some positive number c, so that one of the Dini derivatives of f at x_0 is different from 0. Since one of the Dini derivatives is 0, $f'(x_0)$ fails to exist.

A slight modification in our construction will overcome this difficulty. From each component (a_{nk}, b_{nk}) of G_n we delete a sequence of points that has a_{nk} and b_{nk} as its only limit points. Since E has measure 0, we can choose this sequence so that none of its points belongs to E, and so that the remaining subintervals of (a_{nk}, b_{nk}) are as short as we please. Instead of constructing a mound over (a_{nk}, b_{nk}), we construct a mound over each of the subintervals. If the subintervals are short enough, then for each point x_0 outside of (a_{nk}, b_{nk}) and each point x in (a_{nk}, b_{nk}),

$$|\eta_n(x) - \eta_n(x_0)| < 2^{-n}(x - x_0)^2.$$

Of course, instead of G_2 we must use the intersection of G_2 with what remains after the deletion of the sequences from G_1, and we must make analogous changes in $G_3, G_4, \cdots$.

Since the modification in our construction does not affect properties (i) and (iii) of f, and since it guarantees differentiability of f outside of E, Lemma 2 is now established.

This note was written with support from NSF Grants 1743 and GP 3466.

References

1. P. du Bois-Reymond, Versuch einer Classification der willkürlichen Funktionen reeller Argumente nach ihren Änderungen in den kleinsten Intervallen, J. Reine Angew. Math., 70 (1875) 21–37.

2. K. Weierstrass, Abhandlungen aus der Funktionenlehre, Julius Springer, Berlin, 1886.

3. Z. Zahorski, Punktmengen, in welchen eine stetige Funktion nicht differenzierbar ist, Mat. Sb. N. S., 9 (51) (1941) 487–510 (Russian; German summary).

SELFADJOINT DIFFERENTIAL EXPRESSIONS OF ODD ORDER

A. M. KRALL, Pennsylvania State University

Consider the linear differential operator

$$Ly = Ay^{(2n)} + \sum_{i=0}^{n-1} \binom{2n}{2i+1} C_{2n-2i-1} A^{(2n-2i-1)} y^{(2i+1)},$$

where $C_{2r-1} = r^{-1}(2^{2r}-1)B_{2r}$ with $B_2, B_4, \cdots$, the Bernoulli numbers, $z^{(k)} = d^kz/dx^k$, and A depends only on x. In [1] it is shown that L is selfadjoint. That is, if M is the adjoint of L then $Ly = My$.

Further it was shown that the most general operator satisfying $Ly = My$ is of the form

$$Ly = \sum_{s=0}^{n} \sum_{k=0}^{2s} (-1)^{k+1} \binom{2s}{k} \frac{2^{2s-k+1}-1}{2s-k+1} 2B_{2s-k+1} A_s^{(2s-k)} y^{(k)}.$$

It is sometimes convenient to extend the idea of self-adjointness by saying that the nth order differential operator L is selfadjoint if $Ly = (-1)^n My$. When n is even, this reduces to the previous case. When n is odd, selfadjoint operators now exist. Under the old restrictive definition they did not.

THEOREM. *The differential operator*

$$Ly = Ay^{(2n+1)} + \sum_{j=0}^{n} \binom{2n+1}{2j} C_{2n-2j+1} A^{(2n-2j+1)} y^{(2j)},$$

where $C_{2r-1} = r^{-1}(2^{2r}-1)B_{2r}$ *and* $B_2, B_4, \cdots$, *are the Bernoulli numbers, is self-adjoint.*

Proof. We simplify the adjoint.

$$-My = (Ay)^{(2n+1)} - \sum_{i=0}^{n} \binom{2n+1}{2i} C_{2n-2i+1}(A^{(2n-2i+1)}y)^{(2i)}$$

$$= \sum_{j=0}^{2n+1} \binom{2n+1}{j} A^{(2n+1-j)} y^{(j)} - \sum_{i=0}^{n} \sum_{j=0}^{2i} \binom{2n+1}{2i}\binom{2i}{j} C_{2n-2i+1} A^{(2n+1-j)} y^{(j)}.$$

Since

$$\binom{2n+1}{2i}\binom{2i}{j} = \binom{2n+1}{j}\binom{2n+1-j}{2n-2i+1},$$

by splitting $-My$ into sums of even and odd derivatives of y, using

$$\sum_{i=0}^{n}\sum_{j=0}^{i} = \sum_{j=0}^{n}\sum_{i=j}^{n}, \qquad \sum_{i=0}^{n}\sum_{j=0}^{i-1} = \sum_{j=0}^{n-1}\sum_{i=j+1}^{n}$$

and then letting $k = n - i$, we find

$$-My = Ay^{(2n+1)} + \sum_{j=0}^{n} \binom{2n+1}{2j} A^{(2n-2j+1)} y^{(2j)} \left[1 - \sum_{k=0}^{n-j} \binom{2n-2j+1}{2k+1} C_{2k+1}\right]$$

$$+ \sum_{j=0}^{n-1} \binom{2n+1}{2j+1} A^{(2n-2j)} y^{(2j+1)} \left\{1 - \sum_{k=0}^{n-j-1} \binom{2n-2j}{2k+1} C_{2k+1}\right\}.$$

Theorem 1 of [1] is now applicable. If $n-j+1 = r$, $k = i-1$, the expression in the brackets is seen to be $C_{2n-2j+1}$. With $k = i-1$, the expression in the braces is 0.

COROLLARY. *The most general differential operator satisfying* $Ly = -My$ *is*

$$Ly = \sum_{s=0}^{n} \sum_{k=0}^{2s} (-1)^k \binom{2s+1}{k} \frac{2^{2s-k+2}-1}{2s-k+2} 2B_{2s-k+2} A_s^{(2s-k+1)} y^{(k)}.$$

This follows since $B_1 = -\frac{1}{2}$, $0 = B_3 = B_5 = \cdots$.

Reference

1. H. L. Krall, Selfadjoint differential expressions, this MONTHLY, 67 (1960) 876–878.

SOME REMARKS ON VARIATIONS AND DIFFERENTIALS

M. Z. NASHED, Georgia Institute of Technology

1. Introduction. The treatment of differential calculus for functions of several variables is often dominated in undergraduate courses by computational formulas for differentials, gradients, directional derivatives, etc. The *intrinsic* nature of calculus and the *conceptual* meaning of these notions are seldom brought to light.

The calculus of functions of several variables is in many respects more subtel than the calculus of one real variable. For example, different notions of differentiability can be defined; the mean value theorem and Taylor's formula can be generalized in several ways and the theory of extrema is more involved.

It is generally recognized that the differential and integral calculus of several variables can be best studied in the setting of modern differential geometry, rather than in the traditional setting of real variable theory. For an interesting discussion of the merits of this approach, see [**12**]. On the other hand, an abstract formulation of some aspects of differential calculus can be given using vector spaces and linear operators. The undergraduate student in mathematics today is exposed to these notions in his study of linear algebra and analysis, and can climb up to such levels of formulations.

This approach sheds light on the ideas and arguments of multivariate calculus and the calculus of mapping on normed linear spaces, and unifies many notions and methods in analysis related to integral equations, the calculus of variations and numerical analysis. See, for instance, [**1**, **9**, **11**, **14**, **16**].

The purpose of this exposition is to discuss this approach for mappings whose domain and range are in normed linear spaces over the field of real numbers.

2. Linear and multilinear operators. To make the discussion in the following pages self-contained, we shall review in this section elementary properties of linear and multilinear operators. More details and numerous examples may be found, for instance, in the expository paper by Goffman in [**2**] and in [**3**, Chapter 5].

Throughout this paper, let E be a vector space over the real numbers and let θ denote the zero element in E. A norm on E is a mapping $\|\cdot\|$ which assigns to each element x a real number $\|x\|$ and which satisfies the following axioms:

(i) $\|x\| \geqq 0$; $\|x\| = 0$ only if $x = \theta$,

(ii) $\|\lambda x\| = |\lambda|\ \|x\|$ for any scalar λ,

(iii) $\|x+y\| \leqq \|x\| + \|y\|$.

A normed linear space is a vector space with a norm. Any normed linear space is also a metric space with the distance function $d(x, y) = \|x-y\|$. A complete normed linear space is called a Banach space.

Let E and Y be normed linear spaces. An operator L on E with range in Y is called

(i) additive if $L(x+y)=Lx+Ly$ for all x, y in E,
(ii) homogeneous if $L(\lambda x)=\lambda Lx$ for any scalar λ,
(iii) continuous at x in E if $\|Lx_n-Lx\|\to 0$ as $\|x_n-x\|\to 0$,
(iv) bounded if there exists a nonnegative number M such that $\|Lx\|\leqq M\|x\|$ for all x in E.

L is called *linear* if it is additive and homogeneous. It is well-known that a linear operator is continuous on the whole space if and only if it is continuous at $x=\theta$, and that it is bounded if and only if it is continuous.

If L is a bounded linear operator, then the *norm* of L, denoted by $\|L\|$, is defined as the greatest lower bound of all M satisfying the inequality in (iv). If L and S are bounded linear operators on E to E, then their sum $L+S$ and product LS are again linear operators and

$$\|L+S\| \leqq \|L\| + \|S\| \quad \text{and} \quad \|LS\| \leqq \|L\|\,\|S\|.$$

The set of all bounded linear operators on E to Y forms a Banach space if Y is complete [**3**, p. 102]. Denote this space by $\mathfrak{L}_1$. In particular, the Banach space of all bounded linear transformations from a Banach space E into its field of scalars is called the *dual* space of E and is denoted by E^*. These transformations are called linear functionals over E.

A bilinear operator L_2 on a normed linear space E is a bounded linear operator mapping E into $\mathfrak{L}_1$, i.e. L_2x is in $\mathfrak{L}_1$ and therefore

$$L_2x_1(x_2+x_3) = L_2x_1x_2 + L_2x_1x_3$$

$$L_2(x_1+x_2)x_3 = L_2x_1x_3 + L_2x_2x_3,$$

and

$$\|L_2x_1x_2\| \leqq M\|x_1\|\,\|x_2\|.$$

Similarly we may consider an m-linear operator L_m on a normed linear space E, mapping E into Y, as a bounded linear operator mapping E into $\mathfrak{L}_{m-1}$, where $\mathfrak{L}_{m-1}$ is the space of all $(m-1)$-linear operators in E. More generally, an m-linear operator may be defined on different normed linear spaces.

DEFINITION 1. *A mapping L_m defined on the product space $E_1\times\cdots\times E_m$ of m normed linear spaces E_i, $i=1,\cdots,m$, with range in a normed linear space Y, is called multilinear if it is additive and continuous (hence homogeneous and bounded) in each argument.*

A basic notion associated with multilinear operators is the notion of symmetry.

DEFINITION 2. *A multilinear operator is said to be symmetric if $E_1=E_2=\cdots=E_m$ and $L_m(x_1,\cdots,x_m)$ remains invariant under all permutations of the elements $x_1,\cdots,x_m$.*

3. Bibliographical comments on differential calculus in normed spaces. In building up necessary tools for studying a nonlinear functional in the neighborhood of a fixed element of a normed linear space, it is natural to seek a generalization of either the gradient or the differential as defined in the classical analysis of three-dimensional Euclidean spaces. To include mappings (as well as functionals) it turns out to be more expedient to first generalize the concepts of the differential and the directional derivative. These generalizations were first undertaken by Fréchet [**4**] and Gâteaux [**5**], respectively. Other definitions of differentials were given later by Michal, Zorn, Hyers (see [**8**] for references). Several generalizations of the differential to topological vector spaces, which are not necessarily normed, have been introduced by these and other authors. Such extensions have found useful applications in general differential geometry, dynamics and continuous group theory.

The most useful generalizations of the differential are those which preserve a fundamental idea in calculus, namely the "local" approximation of functions by linear functions. A differential $DF(x; h)$ of a mapping $F: X \to Y$, where X is an open subset of E, E and Y are normed linear spaces, is then a mapping $DF: X \times E \to Y$, where for each x in X, $DF(x; h)$ is assumed—in most definitions—to be linear and continuous in h. The basic difference in the various definitions of a differential is the sense in which $DF(x; h)$ approximates $F(x+h) - F(x)$.

Rolle's theorem does not hold for arbitrary sets in normed linear spaces of dimension greater than one. For such spaces, there are several forms of the mean-value theorem, Taylor's formula and theorem, and each form holds only in a certain sense. The validity of these extensions was established by Graves and Hildebrandt [**6, 7**], who also generalized the implicit function theorem, Kerner [**10**], Kantorovich [**9**], Vainberg [**16**] and others. The proofs used are different in most cases from the proof used in the case of a function of a real variable and as expected a slight strengthening of the hypotheses is required. Kerner considered integrability conditions of abstract vector fields and generalized Stoke's theorem. Rothe [**15**] studied topological properties of gradient mappings, which are generalizations of the idea of a conservative "force field" to abstract vector spaces.

4. Variations. We shall first discuss the Gâteaux variation, which is a generalization of the directional derivative in classical calculus and of the notion of the first variation arising in the calculus of variations.

Let Y be a Banach space and let (t_1, t_2) be an open interval of the real line. The first derivative $\Phi'(t_0)$ of $\Phi: (t_1, t_2) \to Y$, at t_0 in (t_1, t_2) is defined by

$$\Phi'(t_0) = \lim_{t \to t_0} \frac{\Phi(t) - \Phi(t_0)}{t - t_0}$$

if the limit exists, where the limit is taken in the sense of the norm of Y. We note that this derivative is unique and that if Φ has a first derivative at a point t_0, then Φ is continuous at t_0. Higher order derivatives are defined inductively as in classical analysis.

DEFINITION 3. *Let F be a mapping from an open subset X of E into Y, where E and Y are normed linear spaces. Let x_0 be a point in X and h an arbitrary nonzero fixed element in E. Then x_0+th is in X for $|t| \leqq \epsilon(x_0; h)$. Let*

$$\tau = \sup \{\epsilon: \, |t| \leqq \epsilon \Rightarrow x_0 + th \text{ in } X\}.$$

Then $F(x_0+th)$ is defined for $|t|<\tau$. If

$$\frac{d}{dt} F(x_0 + th) \big|_{t=0}$$

exists, it is called the Gâteaux variation (or the weak differential) of F at x_0 with increment h and is denoted by $\delta F(x_0; h)$. If F has a Gâteaux variation, hereafter called G-variation, at every point x in X, then F is said to have a first variation on X.

Similarly, F has an nth variation $\delta^n F(x_0; h)$ at a point x_0 if the function $F(x_0+th)$ has an nth derivative with respect to t at $t=0$.

It follows from the definition, that the first variation is homogeneous in h of degree one, i.e. if $\delta F(x; h)$ exists, then for any scalar λ, $\delta F(x; \lambda h)$ exists and is equal to $\lambda \delta F(x; h)$. Similarly, $\delta^n F(x; \lambda h) = \lambda^n \delta F(x; h)$.

It should be emphasized, however, that the weak differential is not necessarily linear nor continuous in h, as may be seen from the following:

Example 1.

$$f(x_1, x_2) = \frac{x_1 x_2^2}{x_1^2 + x_2^2}, \qquad (x_1, x_2) \neq (0, 0); \qquad f(0, 0) = 0.$$

For each $h=(h_1, h_2)$, the G-variation exists and is equal to $h_1h_2^2(h_1^2+h_2^2)^{-1}$, but the mapping $(h_1, h_2) \to h_1h_2^2(h_1^2+h_2^2)^{-1}$ is not linear in h.

The reader may contrast this remark and the next few results with properties of differentials and variations in *complex* normed spaces, where a different situation prevails [**8**, **11**].

Note that if F has a Gâteaux variation at x_0, then F is continuous in the direction h, i.e.

$$\lim_{t \to 0} \|F(x_0 + th) - F(x_0)\| = 0 \quad (h \text{ is fixed})$$

but is not necessarily continuous at x_0.

Example 2. Let E be the space of all functions $y=y(x)$ which have a continuous first derivative on $[a, b]$. Define a norm on E by

$$\|y\| = \max |y(x)| + \max |y'(x)|,$$

where the maximum is taken over $[a, b]$. Let $f(x, y, z)$ be a function which is defined and has continuous partial derivatives for all finite z and for $a \leqq x \leqq b$, $\Phi_1(x) \leqq y \leqq \Phi_2(x)$ for some prescribed functions Φ_1 and Φ_2. Let

$$J[y] = \int_a^b f(x, y(x), y'(x))\, dx.$$

Then a simple computation shows the G-variation of J at y, corresponding to the increment $h = h(x)$ is

$$\delta J[y; h] = \int_a^b [h(x) f_y(x, y, y') + h'(x) f_{y'}(x, y, y')]\, dx$$

which is the usual first variation.

5. Gâteaux differential. From the above remarks it is clear that the Gâteaux variation does not possess many of the important properties of total differentials for functions of several variables. This motivates the definitions of the Gâteaux and Fréchet differentials, hereafter called G- and F-differentials respectively, which will then enable us to study more effectively a functional or a nonlinear operator in the neighborhood of a fixed element in the space.

DEFINITION 4. *If* $\delta F(x_0; h)$ [Definition 3] *is linear and bounded in* h, *it is called the Gâteaux differential of* F *at* x_0 *with increment* h *and is denoted by* $DF(x_0; h)$.

The G-differential provides in some sense a local approximation property. More precisely, we have

THEOREM 1. *Let* X *be an open subset of* E *and let* F *be a nonlinear operator from* X *to* Y. *A necessary and sufficient condition for* F *to be* G*-differentiable at* x_0 *is that the following representation holds*:

$$F(x_0 + h) - F(x_0) = L(x_0; h) + R(x_0; h) \tag{5.1}$$

for every h *in* E *for which* $x_0 + h$ *is in* X, *where* $L(x_0; h)$ *is linear and continuous in* h *and*

$$\operatorname*{Lim}_{\tau \to 0} \frac{\|R(x_0; \tau h)\|}{\tau} = 0 \qquad \textit{for each } h. \tag{5.2}$$

Proof. We first remark that if such a representation exists, then it is unique. For if another representation exists with L' and R', then

$$\begin{aligned} L(x_0; h) - L'(x_0; h) &= \lim_{\tau \to 0} \tau^{-1}[L(x_0; \tau h) - L'(x_0; \tau h)] \\ &= \lim_{\tau \to 0} \tau^{-1}[R'(x_0; \tau h) - R(x_0; \tau h)] = 0. \end{aligned}$$

Now if the representation (5.1) holds, then

$$\begin{aligned} \left.\frac{dF(x_0 + \tau h)}{d\tau}\right|_{\tau = 0} &= \lim_{\tau \to 0} \tau^{-1}[F(x_0 + \tau h) - F(x_0)] \\ &= L(x_0; h) + \lim_{\tau \to 0} \tau^{-1} R(x_0; \tau h) = L(x_0; h). \end{aligned}$$

Thus the G-variation exists and is linear and continuous in h. Conversely, if the G-differential exists, then

$$\tau^{-1}[F(x_0 + \tau k) - F(x_0)] = DF(x_0; k) + \epsilon(x_0; \tau k),$$

where $\epsilon(x_0; \tau k) \to 0$ as $\tau \to 0$. Letting $\tau k = h$, we get the representation (5.1), where $R(x_0; h) = \tau\epsilon(x_0; h)$ and thus (5.2) holds.

This theorem brings us closer to our objectives and a strengthening of condition (5.2) leads to the definition of the Fréchet differential in Section 6. In the rest of this section, we shall discuss other conditions for a G-variation to be a G-differential.

THEOREM 2. *A necessary and sufficient condition for $\delta F(x_0; h)$ to be linear and continuous in h is that F satisfies the following two conditions:*

(a) *To each h corresponds a $\delta(h)$ such that*

$$|t| \leqq \delta \text{ implies } \|F(x_0 + th) - F(x_0)\| \leqq M\|th\|,$$

where M does not depend on h.

(b) $\overset{2}{\Delta}_{th_1, th_2} F(x_0) = o(t)$ *where*

$$\overset{2}{\Delta}_{h_1, h_2} F(x_0) = F(x_0 + h_1 + h_2) - F(x_0 + h_1) - F(x_0 + h_2) + F(x_0).$$

The proof of the theorem is straight-forward and is given for instance in [**16**, p. 39].

It was noted that $\delta F(x; h)$ is not necessarily linear nor continuous in h or x. It turns out, however, that if $\delta F(x; h)$ is continuous in x at x_0, then it is linear in h. More precisely we state:

THEOREM 3. (See, for instance [**16**, p. 37].) *If F has a G-variation in an open set U such that $\delta F(x; h)$ is continuous in x at some x_0 in U, then $\delta F(x; h)$ is additive in h, i.e. $\delta F(x_0; h+k)$ exists and is equal to $\delta F(x_0; h) + \delta F(x_0; k)$.*

Combining this result with the well-known property of linear operators stated in Section 2, we arrive at

THEOREM 4. *Let the G-variation of the operator F exist in some neighborhood of the point x_0 and let $\delta F(x; h)$ be continuous in x at x_0. Furthermore, assume that $\delta F(x_0; h)$ is continuous in h at $h = \theta$. Then $\delta F(x_0; h)$ is a G-differential.*

6. Fréchet differential.

DEFINITION 5. *The operator F is said to be Fréchet (strongly, totally) differentiable at x_0 if the representation* (5.1) *holds, where $L(x_0; h)$ is linear and continuous in h and moreover*

$$\lim_{h \to \theta} \frac{\|R(x_0; h)\|}{\|h\|} = 0. \tag{6.1}$$

The uniqueness of the Fréchet differential is a special case of the uniqueness of Gâteaux differential, in view of Theorem 5.

We write $L(x_0;\, h) = dF(x_0;\, h) = F'_{x_0}h$ and call it the F-differential of F at x_0 with increment h. The mapping $dF(x_0;\, \cdot) = F'_{x_0}(\cdot)$ which is a bounded linear operator is called the Fréchet *derivative* of F at x_0. It may be noted that the F-derivative is an element in the space $\mathfrak{L}_1$ (see Section 2), while the F-differential $dF(x_0;\, h)$ is an element in Y. This fact is obscured in the calculus of one real variable, where the derivative at a point is defined as a *number*, by the one-to-one correspondence that exists in this case between numbers and linear operators.

It is easy to show that if F is F-differentiable at x_0, then it is continuous at that point; this is not necessarily true, however, if F is G-differentiable. Furthermore, if F is continuous at x_0 then the requirement of continuity of $dF(x_0;\, h)$ in h in Definition 5 is redundant. This follows from the inequality

$$\|dF(x_0;\, h)\| \leqq \|F(x_0 + h) - F(x_0) - dF(x_0;\, h)\| + \|F(x_0 + h) - F(x_0)\|,$$

which shows that $dF(x_0;\, h)$ is continuous at $h = \theta$ and hence continuous everywhere.

REMARK 1. The norm $\|\cdot\|'$ is said to be equivalent to the norm $\|\cdot\|$ if there exist positive numbers m and M such that

$$m\|x\| \leqq \|x\|' \leqq M\|x\|$$

for all x in E. This is clearly an equivalence relation. The definitions of the differentials in the representations (5.1), (5.2), and (6.1) are given in terms of the norms on X and Y. However, it is easy to check that two equivalent norms lead to the same definitions of differentiability, i.e. (5.2) and (6.1) still hold if the norms are replaced by equivalent norms. In the case of a finite-dimensional space all norms are equivalent, so that the differentiability of a mapping F on E into Y is independent of the norms on E and Y, in addition to being independent of the coordinates.

Equivalent norms define the same "topology" so that differentiability depends only on the topologies of X and Y, in infinite-dimensional spaces.

Thus it is possible to extend the definitions of the F- and G-differentials to certain topological vector spaces. It may also be observed in view of (5.2) that G-differentiability is meaningful if E is a linear space and Y a topological linear space.

Example 3. Let $F\colon E^n \to E^m$, where E^n is the Euclidean n-space. That is, $y_i = F_i(x_1, \cdots, x_n)$, $i = 1, 2, \cdots, m$. Assume that F has an F-differential at $a = (a_1, \cdots, a_n)$. Then it is not hard to show that the partial derivatives $\partial F_i/\partial x_j$ exist at a, and that the F-derivative is the linear transformation whose matrix is $[\partial F_i/\partial x_j]$. Conversely, assume that there exists an $r > 0$ such that for x in $\|x - a\| \leqq r$, $\partial F_i/\partial x_j$ exist, and are continuous at $x = a$, then F has an F-differential at $x = a$. See, for instance, *Advanced Calculus*, by R. C. Buck.

Example 4. Let $K(s, t)$ be a continuous real function for $0 \leqq s, t \leqq 1$, and assume that K is symmetric, i.e. $K(s, t) = K(t, s)$. The functional

$$J[x] = \int_0^1 x^2(t)dt - \lambda \int_0^1 \int_0^1 K(s, t)x(s)x(t)dsdt$$

is defined on the space of all continuous real functions on $[0, 1]$ with norm $\|x\| = \max_{0 \leq t \leq 1} |x(t)|$. From a simple computation it follows that

$$\frac{d}{d\tau} J[x + \tau h] \Big|_{\tau=0} = 2 \int_0^1 x(t)h(t)\, dt - 2\lambda \int_0^1 \int_0^1 K(s, t)x(s)h(t)\, dtds$$

which is linear and continuous in h. Condition (6.1) is also satisfied, so that the last expression is the F-differential of J.

7. Gradients. The definitions and properties given in the preceding sections are for mappings between normed linear spaces. Thus, they hold in particular for a functional f defined on a subset X of a normed linear space E and mapping each x in X into a real number $f(x)$. By Definitions 4 and 5, if f is differentiable on X then there exists a mapping df on $X \times E$ into the reals, which is linear and continuous in h in E. Another way of looking at df is to consider $df(x_0; h)$ as a continuous linear functional for each fixed x_0, and denote it by $f'_{x_0} h$. Then f'_{x_0} is an element of the dual space E^* of E. As x_0 varies over X, a mapping $f'_x(\cdot): X \to E^*$ is thus obtained, which Rothe [**15**] called the *gradient* mapping of f.

For example if f is a differentiable function of three real variables, then the differential of f at $\vec{x} = (x_1, x_2, x_3)$, with increment $\vec{h} = (h_1, h_2, h_3)$ is given by

$$(7.1) \qquad df(\vec{x}; \vec{h}) = \sum_{i=1}^{3} \frac{\partial f}{\partial x_i} h_i = \vec{\Gamma}(x) \cdot \vec{h}, \quad \text{where} \quad \vec{\Gamma}(x) = \operatorname{grad} f(x).$$

Thus $\vec{\Gamma}(x)$ assigns to each x in E a continuous linear functional df. In this example df is the inner product of grad f and h.

Similarly, if E is a complete inner product space H, ([2, p. 103] or [**3**, p. 111]) then the gradient mapping may be considered as a mapping from H into itself since the dual space may be indentified with H. Furthermore, $df(x_0; h)$, being a continuous linear functional in h, can be uniquely represented as an inner product [**3**, p. 117], i.e. there exists a unique $\Gamma(x_0)$ in H such that

$$(7.2) \qquad df(x_0; h) = (\Gamma(x_0), h),$$

where the parentheses denote inner product. $\Gamma(x_0)$, defined by (7.2) is called the gradient of the functional f at x_0 and is denoted by grad $f(x_0)$.

For instance it follows from Example 4, under the usual inner product $\int_a^b x(t)y(t)dt$, that

$$\tfrac{1}{2} \operatorname{grad} J[x] = x(t) - \lambda \int_0^1 K(s, t)x(s)\, ds.$$

The concept of gradient in abstract spaces was first introduced by M. Golomb in his study of nonlinear integral equations.

If in (7.2), we use the Gâteaux differential, then we obtain the definition of the *weak* gradient grad_w of f at x_0, i.e.

$$Df(x_0; h) = (\text{grad}_w f(x_0), h). \tag{7.3}$$

REMARK 2. The gradient depends on the inner product. For instance if P is a positive operator, i.e. $(Px, x) > 0$ unless $x = \theta$, then we may define a *new* inner product $[x, y]$ by

$$[x, y] = (Px, y). \tag{7.4}$$

Let $\text{grad}\, f$ and $\text{grad}^{\#} f$ denote the gradient of f with respect to the original and new inner products respectively. Then by definition of gradient,

$$df(x; h) = [\text{grad}^{\#} f, h] = (\text{grad}\, f, h). \tag{7.5}$$

But from (7.4)

$$[\text{grad}^{\#} f, h] = (P\, \text{grad}\, f, h). \tag{7.6}$$

From (7.5) and (7.6) we get for all h, $(\text{grad}\, f, h) = (P\, \text{grad}^{\#} f, h)$. Hence,

$$\text{grad}\, f = P\, \text{grad}^{\#} f.$$

For example, if in E^3 we define "distance" by

$$d(x, y) = \left\{ \sum_{i,j=1}^{n} p_{ij}(y_j - x_j)(y_i - x_i) \right\}^{1/2} = (P(x - y), x - y)^{1/2},$$

where $P = [p_{ij}]$ is a positive definite matrix, then the gradient of f with respect to this metric is related to the gradient of f with respect to the usual Euclidean distance by

$$\text{grad}^{\#} f = [p_{ij}]^{-1}\, \text{grad}\, f.$$

8. Implication relationships between F- and G-differentiability. The only difference between the Fréchet and Gâteaux differentials is in the relations (5.2) and (6.1). We now show that (6.1) implies (5.2) but not conversely. This result is included in the following interesting characterization of the F-differential.

THEOREM 5. *The operator F is F-differentiable at x_0 if and only if the representation* (5.1) *holds, where $L(x_0; h)$ is continuous and linear in h and*

$$\lim_{\tau \to 0} \tau^{-1} \| R(x_0; \tau h) \| = 0 \tag{8.1}$$

uniformly with respect to h on each set $\|h\| = constant$.

Proof. Without any loss of generality, we may prove this for the set $\|h\| = 1$. If F is Fréchet differentiable at x_0, then

$$\lim_{||h||\to 0} \frac{||R(x_0;\, h)||}{||h||} = 0.$$

Letting $h=\tau k$, where k has a unit norm, we get $\lim_{\tau\to 0}\tau^{-1}||R(x_0;\, \tau k)||=0$ uniformly on $||k||=1$.

Conversely, if (8.1) holds uniformly on each bounded set, then, in view of Theorem 1, F has a G-differential $DF(x_0;h)$ at x_0. Thus for any given $\epsilon>0$, there exists a $\delta>0$ such that

$$||\tau^{-1}[F(x_0+\tau h) - F(x_0)] - DF(x_0;\, h)|| < \epsilon,$$

whenever $|\tau| < \delta$. That is,

$$F(x_0 + \tau h) - F(x_0) = DF(x_0; \tau h) + R(x_0; \tau h),$$

where for $|\tau| < \delta$,

$$\frac{||R(x_0;\, \tau h)||}{||\tau h||} < \epsilon, \qquad (||h|| = 1).$$

Letting $k=\tau h$, we get $F(x_0+k)-F(x_0)=DF(x_0;\, k)+R(x_0;\, k)$, where

$$\lim_{k\to\theta} \frac{||R(x_0;\, k)||}{||k||} = 0.$$

Hence, $DF(x_0;\, k)=dF(x_0;\, k)$.

Thus if F is F-differentiable at x_0, then F is G-differentiable (and consequently it has a G-variation) at x_0. Furthermore

$$dF(x_0;\, h) = DF(x_0;\, h) = \delta F(x_0;\, h).$$

The converse holds if F is a function of one real variable, but does not necessarily hold in higher dimensions as may be seen from the following:

Example 5. Let $x=(x_1,\, x_2)$ where x_1 and x_2 are real variables and consider

$$f(x) = \frac{x_1}{x_2}(x_1^2 + x_2^2), \qquad x_2 \neq 0; \qquad f(x_1,\, 0) = 0,$$

and let $||x||=(|x_1|^2+|x_2|^2)^{1/2}$. Then f has a G-variation at $x=(0,\, 0)$, which is (trivially) continuous and linear in h. In this case,

$$R(0;h) = \frac{h_1}{h_2}(h_1^2 + h_2^2), \quad h_2 \neq 0; \qquad\qquad R(0,\, h) = 0 \text{ if } h_2 = 0,$$

and hence (5.2) holds. However, the F-differential does not exist at (0, 0). For if we let $h_n=(n^{-1/2},\, n^{-1})$ then $h_n\to\theta$ while

$$\frac{||R(0;\, h_n)||}{||h_n||} = \sqrt{1 + 1/n} \to 1, \qquad\qquad \text{as } n \to \infty.$$

REMARK 3. Neither the G-differential $Df(x_0; h)$ nor the F-differential $df(x_0; h)$ of a functional f is required to be continuous in x at x_0; and hence neither the gradient nor the weak gradient of f. It can be shown, however, that if $\operatorname{grad}_w f(x)$ exists and is continuous in x on an open set Ω, then it coincides with $\operatorname{grad} f(x)$. This then implies that f is F-differentiable and $df(x; h) = Df(x; h)$.

REMARK 4. The requirement of continuity of grad f is related to the notion of *uniform* Fréchet differentials. F is said to have a *locally uniform* F-differential $dF(x; h)$ on an open set Ω if F has an F-differential on Ω and the remainder $R(x_0; h)$ is locally uniformly bounded, i.e. for each $\epsilon > 0$ and an arbitrary x_0 in Ω, there exists a $\delta(x_0; \epsilon)$ and $\eta(x_0; \epsilon)$ such that

$$\|R(x; h)\| \leqq \epsilon\|h\| \quad \text{if} \quad \|h\| \leqq \delta \quad \text{and} \quad \|x - x_0\| \leqq \eta.$$

It turns out that a necessary and sufficient condition for grad $f(x)$ to be continuous in the sphere S: $\|x\| < a$ is that $df(x; h)$ have a locally uniform remainder and grad $f(x)$ be locally bounded. We leave the discussion of the interesting implications of these remarks and proofs in the case of a function of several real variables to the reader, (see Example 3).

9. Higher order differentials. The first order differential $dF(x; h)$ is a function of two variables. Thus, several notions for a second order differential may be defined. The most natural notion, which we will discuss here, is based on the observation that $F_x'(\cdot) = dF(x; \cdot)$ which is an element of the space $\mathfrak{L}_1$, is also an *operator* sending X into the space $\mathfrak{L}_1$. If this operator is F-differentiable, its derivative is called the second order Fréchet derivative of F and is denoted by $F_x''(\cdot, \cdot)$. Thus the second order derivative is an element of the space $\mathfrak{L}_2$ of all continuous linear operators from E into $\mathfrak{L}_1$, i.e. it is a bilinear operator from E to Y; it also has the representation

$$dF(x + k; h) - dF(x; h) = F_x'' hk + R(x; h, k),$$

where

$$\lim_{k \to 0} \frac{\|R(x; h, k)\|}{\|k\|} = 0.$$

$F_x'' hk = d^2F(x; h, k)$ is called the second order Fréchet differential of F.

The F-differential of the nth order may be defined inductively as follows:

DEFINITION 6. *Let E and Y be normed linear spaces over the field of real numbers and X be an open subset of E. Suppose that for some integer $m \geqq 2$, the m-th order F-differential $d^mF(x_0; h_1 \cdots h_m)$ of the mapping F: $X \to Y$, has been defined for all $(m+1)$-tuples $(x_0, h_1, \cdots, h_m)$ of elements of E such that $x_0 + \sum_{i=1}^{m} h_i$ is in X. Then F is said to have an F-differential of order $m+1$, if for all $(m+2)$-tuples of elements $(x_0, h_1, \cdots, h_{m+1})$ of E such that $x_0 + \sum_{i=1}^{m+1} h_i$ is in X, the following representation holds*:

$$d^mF(x_0 + h_{m+1}; h_1, \cdots, h_m) - d^mF(x_0; h_1, \cdots, h_m)$$
$$= d^{m+1}F(x_0; h_1, \cdots, h_m, h_{m+1}) + R(x_0; h_1, \cdots, h_{m+1}),$$

where the mapping $d^mF(x; h_1, \cdots, h_{m+1})$ *is linear and continuous in* h_{m+1} *and*

$$\lim_{||h_{m+1}||\to 0} \|h_{m+1}\|^{-1}\|R(x_0; h_1, \cdots, h_m, h_{m+1})\| = 0.$$

If such a representation exists, it is unique and $d^{m+1}F(x_0; h_1, \cdots, h_{m+1})$ is called the $(m+1)$-th F-differential of F at x_0. The operator $d^{m+1}F(x_0; \cdots)$ is called the $(m+1)$-th Fréchet derivative. An inductive argument shows that $d^{m+1}F(x_0; h_1, \cdots, h_{m+1})$ is $(m+1)$-linear in $h_1, \cdots, h_{m+1}$ (Definition 1). Symmetry of the mth F-differential in $h_1, \cdots, h_m$ may therefore be defined (Definition 2). An interesting result, which is a generalization of the sufficient condition that makes immaterial the order of mixed partial differentiation for real functions of several variables, may be stated as follows: a sufficient condition for $d^mF(x; h_1, \cdots, h_m)$ to be symmetric in $h_1, \cdots, h_m$ at $x=x_0$ is that $d^mF(x; h_1, \cdots, h_m)$ be continuous in x for all x in some neighborhood of x_0 (Theorem 8 in [**6**]).

The Gâteaux differential of order m may be defined similarly; implication relationships similar to those established in Section 8 may be stated. For example, if F has an F-differential of the mth order at x_0, then the mth order variation of F at that point exists and moreover

$$d^mF(x_0; h_1, \cdots, h_m) = \frac{\partial^m}{\partial t_1 \cdots \partial t_m} F\left(x_0 + \sum_{i=1}^{m} t_i h_i\right)\Bigg|_{t_1=\cdots=t_m=0}.$$

In particular, if $h_1 = \cdots = h_m = h$, then

$$d^mF(x_0; h, \cdots, h) = \frac{d^m}{dt^m} F(x_0 + th)\Bigg|_{t=0}.$$

10. Remarks.

A. It should be noted that the notions presented in the preceding sections are *coordinate free* and that the derivative is defined in an invariant form as a linear transformation. This approach is useful in applied mathematics for formulating simultaneous algebraic equations, integral equations and boundary-value problems, etc., as operator equations, and for approximate and iterative methods for solving these equations. See for instance the interesting expository paper on Newton's method and variations by R. H. Moore in [**1**] and [**9, 14, 16**]. The compactness of notation that results from this approach and the abstractness of the notions are assets to the conceptual framework, unifying diverse situations in analysis and approximation theory.

B. Some of the recent books in advanced calculus and real variables, treat differentials of mapping from regions in E^n to E^m in the spirit of linear transformations. We refer specifically to the outstanding books on *Advanced Calculus*,

by T. M. Apostol; R. C. Buck; W. Maak; H. K. Nickerson, D. C. Spencer, and N. E. Steenrod; and to *Principles of Mathematical Analysis* (2nd edition) by W. Rudin, *The Elements of Real Analysis* by R. G. Bartle and [**3**]. An elementary introduction is also given in *Calculus of Vector Functions* by R. H. Crowell and R. E. Williamson. See also the recent books by W. H. Fleming and C. Goffman on functions of several variables.

C. Differentiation rules, chain rules, mean value theorems, Taylor's formula, etc., can be developed for F- and G-differentials as in classical calculus. Partial F- and G-differentials can also be defined paralleling the classical theory.

11. Some open questions. One mathematician remarked that a colloquium lecture in mathematics should include at least one proof and one open problem. We assume, without further discussion, that this also holds for an expository article. We conclude therefore by mentioning some open questions which can be stated within the framework of this paper.

A. The derivative $dF(x;\ \cdot)$ is not necessarily continuous in x. The question then arises as when we can approximate a nonlinear operator by a continuous Fréchet derivative or, more generally, by another *nonlinear* map with a continuous Fréchet derivative, in the case of infinite-dimensional spaces.

B. It is known that a real function of a real variable which satisfies a Lipschitz condition is differentiable almost everywhere. This follows from the fact that if $|f(x)-f(y)| \leqq M|x-y|$ for all x and y in $[a, b]$, then

$$\sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)| \leqq M \sum_{i=1}^{n-1} |x_{i+1} - x_i|,$$

for any partition $a \leqq x_1 < x_2 < \cdots < x_n \leqq b$. Thus f is of bounded variation, and hence being the difference of two monotonic functions is differentiable almost everywhere. The questions that arise are:

(i) Does an operator which is Lipschitz continuous, i.e.

$$\|F(x) - F(y)\| \leqq M\|x - y\| \quad \text{for all } x \text{ and } y,$$

have any G- or F-differentiability properties almost everywhere?

(ii) What additional hypotheses are sufficient to imply G- or F-differentiability for a Lipschitz continuous operator?

C. A functional defined on a linear space E (or on a convex subset of E) is said to be *convex* if for all x and y in the domain of f and for $0 \leqq a \leqq 1$, $f[ax+(1-a)y] \leqq af(x)+(1-a)f(y)$. A convex functional defined on an open subset has a one-sided G-variation, i.e.

$$\lim_{t\to 0^+} t^{-1}[f(x + th) - f(x)]$$

exists, $t>0$. Furthermore, if f is continuous and convex on the real interval $[a, b]$, then f has a right-hand and left-hand derivative at every point and the subset on which f' does not exist is countable. See, for instance, pages 195–196 in *Analysis* by E. Hille.

We ask the same questions as in **B** for convex functionals, i.e., what hypotheses imply G- or F-differentiability of a convex functional, which would not imply the same for arbitrary functionals?

A similar question can be posed for monotone operators (see the paper by Dolph and Minty in [1]).

References

1. P. M. Anselone, editor, Nonlinear integral equations, Univ. of Wisconsin Press, Madison, 1964.

2. R. C. Buck, editor, Studies in modern analysis, Vol. 1, MAA Studies, 1962.

3. J. Dieudonné, Foundations of modern analysis, Academic Press, New York, 1960.

4. M. Fréchet, La notion de différentielle dans l'analyse générale, Ann. Sci. École Norm. Sup., 42 (1925) 293–323.

5. R. Gâteaux, Sur les fonctionnelles continues et les fonctionnelles analytiques, Bull. Soc. Math. France, 50 (1922) 1–21.

6. L. M. Graves, Riemann integration and Taylor's theorem in general analysis, Trans. Amer. Math. Soc., 29 (1927) 163–177.

7. T. H. Hildebrandt and L. M. Graves, Implicit functions and their differentials in general analysis, Trans Amer. Math. Soc., 29 (1927) 127–153.

8. E. Hille and R. S. Phillips, Functional analysis and semigroups, Amer. Math. Soc. Coll. Publ., Providence, 1957, (see Chapter 26).

9. L. V. Kantorovich and G. P. Akhilov, Functional Analysis in Normed Spaces, Pergamon Press, New York, 1964.

10. M. Kerner, Die Differentiale in der allgemeinen Analysis, Ann. of Math., 34 (1933) 546–572.

11. A. D. Michal, Le calcul différentiel dans les espaces de Banach, vol. 1, Gauthier-Villars, Paris, 1958.

12. M. E. Munroe, Bringing calculus up-to-date, this MONTHLY, 65 (1958) 81–90.

13. M. E. Munroe, Modern Multidimensional Calculus, Addison-Wesley, Reading, Mass., 1963.

14. M. Z. Nashed, On general iterative methods for the solutions of a class of nonlinear operator equations, Math. Comp., 19 (1965) 14–24.

15. E. H. Rothe, Gradient mappings, Bull. Amer. Math. Soc., 59 (1953) 5–19.

16. M. M. Vainberg, Variational Methods for the Study of Nonlinear Operators, Holden-Day, San Francisco, 1964.

17. A. Bronsted and R. T. Rochafellar, On the subdifferentiability of convex functions, Proc. Amer. Math. Soc., 16 (1965) 605–611.

18. E. B. Leach, Differentiability of sub-convex functionals on Banach spaces, Amer. Math. Soc. Notices, 625–164, 1965.

19. T. Rado and P. V. Reichelderfer, Continuous Transformations in Analysis, Springer Verlag, Berlin, 1955.

A FINITE DIFFERENCE EXTENSION OF THE LAW OF THE MEAN

G. E. MATTHEWS, State University of New York at Albany

A lemma of S. Bernstein [1], extending Rolle's Theorem to the calculus of finite differences, states that if f is a continuous real function on $[a, b]$ which vanishes at the endpoints, then the difference $\Delta_h f$, defined by $\Delta_h f(x) = f(x+h) - f(x)$, has a zero in the interval, provided h is sufficiently small. R. J. Levit has recently shown [2] that the best possible bound on h depending only on the number n of changes in sign of f in (a, b) is

$$H_n = \frac{(b - a)}{[(n + 3)/2]},$$

where the brackets denote the greatest integer function.

R. J. Levit's re-statement of Bernstein's Lemma may be expressed as follows: If a real function f is continuous on $[a, b]$ and vanishes at the endpoints and changes sign exactly n times in (a, b) and if $0 < h \leqq H_n = (b-a)/[(n+3)/2]$, then for any given h which satisfies the hypothesis there exists an r, where $a \leqq r \leqq b - h$, such that $\Delta_h f(r)/h = 0$. This is clearly seen to be a finite difference extension of Rolle's Theorem with the property of differentiability of f replaced by a bound on h and with the first derivative replaced by the quotient $\Delta_h f/h$.

The purpose of this note is to extend the Law of the Mean to the calculus of finite differences in the same way with the same bound H_n on h. For a generalized Law of the Mean for finite differences, the reader is referred to a paper by D. Raikov [3], which applies Bernstein's Lemma without such a bound on h.

The Law of the Mean states that if a real function f is continuous on $[a, b]$ and differentiable on (a, b), then there exists an r, where $a < r < b$, such that $f(b) - f(a) = (b-a)f'(r)$. The Law of the Mean may be extended to the calculus of finite differences as in the following new theorem:

THEOREM. *If a real function f is continuous on $[a, b]$ and if the graph of f crosses the secant line joining the points $(a, f(a))$ and $(b, f(b))$ exactly n times in (a, b) and if*

$$H_n = (b - a)/[(n + 3)/2],$$

then corresponding to any positive $h \leqq H_n$ there exists an r, where $a \leqq r \leqq b - h$, such that

$$f(b) - f(a) = (b - a)\Delta_h f(r)/h.$$

Proof. Suppose that the secant line has the equation $y = mx + c$; then consider the auxiliary function ϕ, defined by $\phi(x) = f(x) - mx - c$ for x in $[a, b]$. Since f is continuous on $[a, b]$, then ϕ is continuous on $[a, b]$. Clearly, $\phi(a) = \phi(b) = 0$. If f crosses the line $y = mx + c$ exactly n times in (a, b), then ϕ changes sign (crosses the line $y = 0$) exactly n times in (a, b). Thus H_n has the same

value for both the given function f and the auxiliary function ϕ. It is seen that ϕ satisfies the hypothesis of Levit's Theorem.

Hence, for any positive $h \leqq H_n$, there exists an r, where $a \leqq r \leqq b-h$, such that $\Delta_h\phi(r)/h=0$. Expressing this result in terms of the original function f yields $(\Delta_h f(r)-mh)/h=0$. Thus

$$\Delta_h f(r)/h = m = (f(b) - f(a))/(b - a).$$

The conclusion of the new theorem follows immediately: Corresponding to any positive $h \leqq H_n$, there exists an r, where $a \leqq r \leqq b-h$, such that

$$f(b) - f(a) = (b - a)\Delta_h f(r)/h.$$

Since the preceding theorem implies the existence of chords which are parallel to the secant line connecting the endpoints of the graph of the function, we seek a bound on the length of these chords.

COROLLARY. *If a real function f is continuous on $[a, b]$ and if the graph of f crosses the secant line joining the points $(a, f(a))$ and $(b, f(b))$ exactly n times in (a, b), and if*

$$L = H_n\sqrt{1 + m^2},$$

where m is the slope of the secant line, then corresponding to any positive $l \leqq L$, the graph has a chord in $[a, b]$ of length l parallel to the secant line.

Proof. From the Pythagorean Theorem it can easily be shown that a chord with slope m has a length $l=h\sqrt{1+m^2}$, where h is the length of the projection of the chord on the X-axis. Now, if $l \leqq L$, then $h \leqq H_n$, and the hypothesis of the theorem is satisfied. And the theorem implies the existence of the desired chords for the bound L. Clearly, L is the best possible bound on l as it depends only on m (which is a constant for any given function) and on H_n (which was proved best possible in [2]).

In order to guarantee the existence of an r, and hence a chord of length l, in the open interval (a, b), we must replace H_n with a more restrictive bound $H_n^*=(b-a)/[(n+5)/2]$. (See [2] page 29.) We leave the re-statements and proofs as an exercise for the reader.

References

1. S. Bernstein, Sur la définition et les propriétés des fonctions analytiques d'une variable réelle, Math. Ann., 75 (1914) 452.

2. R. J. Levit, The finite difference extension of Rolle's theorem, this MONTHLY, 70 (1963) 26–30.

3. D. Raikov, Sur les différences et sur les dérivées, Rec. Math. [Mat. Sbornik] N.S., 7(49) (1940) 379–384.

ON THE SPECTRUM OF A LAURENT FORM

M. EISEN, University of Pittsburgh AND H. GINDLER, San Diego State College

1. Introduction. In this paper we study the spectrum of a certain class of linear transformations on l_p (Def. 1) called *Laurent forms* (Def. 2) and show how the results of this theory can be used to give very short and simple proofs of the famous Wiener and Lévy theorems concerning absolutely convergent Fourier series [**15**, pp. 245–7].

In Section 2 we give some preliminary definitions. The operational calculus for bounded linear transformations is introduced in Section 3 and applied to obtain the spectrum of a regular L-form (Def. 5). In Section 4 the fine structure of the spectrum is studied and the spectrum of an arbitrary L-form is obtained in Section 5. This result is used to prove Wiener's theorem in Section 6.

This paper is partially expository since it is written with the objective of making this information accessible to persons who have had only a brief introduction to Functional Analysis. We believe that L-forms could be used very effectively for problems and examples in a graduate course which includes spectral theory for infinite dimensional normed linear spaces.

2. Preliminary notions. In this section we introduce some definitions, facts, and notations which will be used later on.

DEFINITION 1. *The normed linear space l_p, where $1 \leqq p < \infty$, is defined to consist of all doubly infinite sequences $x = \{x_n\}$ of complex numbers x such that $(\sum_{n=-\infty}^{\infty} |x_n|^p)^{1/p} < \infty$. The norm in l_p is defined by*

$$\|x\| = \left(\sum_{n=-\infty}^{\infty} |x_n|^p \right)^{1/p}. \tag{1}$$

The space l_∞ consists of all bounded sequences; the norm in l_∞ is defined to be

$$\|x\| = \sup |x_n|, \tag{2}$$

where the supremum is taken over the set of all integers.

We recall some well-known facts about l_p spaces.

LEMMA 1. *The l_p spaces $(1 \leqq p \leqq \infty)$ are Banach spaces.*

LEMMA 2. *The dual space, the space of all continuous linear functionals, of l_p is congruent to l_q, where $1/p + 1/q = 1$, if $1 \leqq p < \infty$.*

The above lemma is not true for $p = \infty$.

DEFINITION 2. *An L-form (Laurent form) is a doubly infinite matrix (a_{ij}) where $a_{ij} = a_{i-j}$, for $i, j = 0, \pm 1, \pm 2, \cdots$, are complex numbers such that $\sum_{n=-\infty}^{\infty} |a_n| < \infty$.*

More specifically, an L-form L is represented as follows:

$$L = \left| \begin{array}{cccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdots a_0 & a_1 & a_2 & a_3 & a_4 \cdots \\ \cdots a_{-1} & a_0 & a_1 & a_2 & a_3 \cdots \\ \cdots a_{-2} & a_{-1} & a_0 & a_1 & a_2 \cdots \\ \cdots a_{-3} & a_{-2} & a_{-1} & a_0 & a_1 \cdots \\ \cdots a_{-4} & a_{-3} & a_{-2} & a_{-1} & a_0 \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \right| \tag{3}$$

An important role in subsequent developments will be played by the special L-form T obtained when $a_1 = 1$ and $a_n = 0$ if $n \neq 1$.

By indulging in the usual matrix multiplication one sees that an L-form assigns to each doubly infinite bounded sequence of complex numbers another sequence and, just as in the finite dimensional case, this correspondence determines a linear transformation. In fact, the matrix multiplication is merely the convolution of an element of l_1 with an element of l_p.

Example. Let x be the column vector $x = (\cdots, x_{-1}, x_{-2}, x_0, x_1, x_2, \cdots)$ and T be the matrix described in the preceding paragraph. If $y = Tx$, then $y_i = x_{i+1}$ $(i = 0, \pm 1, \pm 2, \pm 3, \cdots)$—that is, T is a shift operator.

DEFINITION 3. *If X is any Banach space we denote by $[X]$ the space of all continuous linear transformations defined on all of X, with range X. If $A \in [X]$ we define the norm of A by*

$$\|A\| = \sup \|Ax\|, \tag{4}$$

where the supremum is to be taken over all x in X with $\|x\| = 1$.

It is known that $\|A\|$ is finite for A in $[X]$ and that $[X]$ is a Banach space (in fact a Banach algebra). The topology on $[X]$ determined by (4) is called the uniform operator topology.

The following lemma is a consequence of the transformation law for T, given in the example, and the above definition.

LEMMA 3. *$T \in [l_p]$ if $1 \leqq p \leqq \infty$ and $\|T\| = 1$.*

Subsequently, we will prove that any L-form belongs to $[l_p]$.

We shall need the notion of the spectrum of an L-form.

DEFINITION 4. *If L is an L-form, then the resolvent set $\rho(L)$ is the set of complex numbers λ such that $L - \lambda I$ is 1:1 onto l_p and $(L - \lambda I)^{-1} \in [l_p]$. Here I denotes the identity matrix. The spectrum $\sigma(L)$ is the set of complex numbers λ such that $\lambda \notin \rho(L)$. Frequently we shall write $L - \lambda$ instead of $L - \lambda I$.*

The above definition is not the usual one, but it is equivalent to the standard definition of the spectrum for the cases we will be considering. In fact, without loss of generality, $\rho(A)$ and $\sigma(A)$ can be defined as above for A in $[X]$ where X is a Banach space.

Note that $\lambda \in \rho(L)$ if and only if the equation $y=(L-\lambda I)x$ has a unique solution x in l_p for each y in l_p and x is a continuous function of y. Technically $\rho(A)$ will depend on p as well as A. However, for L-forms we shall see that $\rho(L)$ does not vary with p.

To illustrate the above concepts we shall find $\rho(T)$ and $\sigma(T)$. The following theorems will be useful for this purpose.

THEOREM 1. *Let X be a Banach space, λ a complex number, and $A \in [X]$. If $|\lambda| > \|A\|$ then $(A-\lambda I)^{-1} \in [X]$.*

For a proof see Taylor [**12**, Theorem 5.2 A, p. 260].

THEOREM 2. *Let T be the L-form given by the matrix (a_{ij}) where $a_{ij}=1$ if $j-i=1$ and $a_{ij}=0$ otherwise. Then, if λ denotes a complex number*

$$\sigma(T) = (\lambda\colon |\lambda| = 1), \tag{5}$$

and

$$\rho(T) = (\lambda\colon |\lambda| \neq 1). \tag{6}$$

Proof. Since $\|T\|=1$ (cf. Lemma 2) and the $l_p (1 \leqq p \leqq \infty)$ are Banach spaces (cf. Lemma 1) it follows by Theorem 1 that $\lambda \in \rho(T)$ if $|\lambda| > 1$.

Let $y=(\lambda I - T)x$, then

$$y_i = \lambda x_i - x_{i+1} \qquad \text{for } i = 0, \pm 1, \pm 2, \cdots. \tag{7}$$

Solving (7) recursively we obtain

$$x_k = \lambda^k x_0 - (\lambda^{k-1}y_0 + \lambda^{k-2}y_1 + \cdots + \lambda y_{k-2} + y_{k-1}) \qquad \text{for } k = 0, 1, 2, \cdots, \tag{8}$$

$$x_{-k} = \lambda^{-k}x_0 + (\lambda^{-k}y_{-1} + \lambda^{-(k-1)}y_{-2} + \cdots + \lambda^{-1}y_{-k}) \qquad \text{for } k = 1, 2, 3, \cdots. \tag{9}$$

Now if $|\lambda|=1$, then setting $y=0$ for all i gives that $\{\lambda^k x_0\}$ is an eigenvector in l_∞ where x_0 is an arbitrary complex number. We conclude that $\lambda \in \sigma(T)$ if $|\lambda|=1$ and $p=\infty$.

If $|\lambda|=1$ and if $x \in l_p$, with $1 \leqq p < \infty$, then $x_k \to 0$ as $k \to \pm\infty$. Letting $k \to \infty$ in (8) and (9) and equating the resulting expressions for x_0 we obtain

$$\sum \lambda^{-i} y_i = 0. \tag{10}$$

Setting $\lambda^{-i} y_i = z_i$, equation (10) becomes

$$\sum_{-\infty}^{\infty} z_i = 0. \tag{11}$$

Since (11) cannot be satisfied by every sequence $\{z_i\}$ in l_p we can find a sequence

$\{y_i\}$ in l_p (since $|\lambda| = 1$) which cannot satisfy (7). We conclude that $\lambda I - T$ cannot be onto l_p. Thus $|\lambda| = 1$ implies $\lambda \in \sigma(T)$.

To complete the proof of the theorem we need only show that $\lambda \in \rho(T)$ if $|\lambda| < 1$. We begin by considering the point $\lambda = 0$. From (7) it is obvious that the range of $-T$ is all of l_p. Further $-T^{-1}$ is a shift operator, in fact an L-form with $a_{-1} = 1$ and $a_i = 0$ if $i \neq -1$, and so $\|-T^{-1}\| = 1$. Boundedness of a linear operator, however, is equivalent to continuity and so $0 \in \rho(T)$. Now let $0 < |\lambda| < 1$, by Theorem 1 it follows just as before that $\lambda^{-1} \in \rho(T^{-1})$ and we have $(\lambda - T)^{-1} = \lambda^{-1}T^{-1}(T^{-1} - \lambda^{-1})^{-1} \in [l_p]$ and so $\lambda \in \rho(T)$.

An interesting feature of the above proof is that λ can be in the spectrum of T without being an eigenvalue (recall the discussion for $|\lambda| = 1$ and $1 \leq p < \infty$).

A direct proof that $(\lambda - T)^{-1} \in [X]$ can be given by actually finding an analytic expression for this operator by using (7). We leave it to the reader to verify that $(\lambda - T)^{-1}$ as given by equations (19) and (20) is correct.

We postpone this study of the fine structure of the spectrum until Section 4 in order to minimize the number of definitions needed at this stage.

3. The operational calculus. Throughout this section let X be a Banach space and suppose that $A \in [X]$ with $\|A\|$ as defined in equation (4).

If $\lambda \in \rho(A)$ with $A \in [X]$, then by definition $A - \lambda I$ maps X, 1:1 onto X. It is known that in this case $(\lambda I - A)^{-1} \in [X]$. We denote by R_λ the operator valued function which assigns to each $\lambda \in \rho(A)$ the operator $(\lambda I - A)^{-1}$; thus for all $\lambda \in \rho(A)$

$$R_\lambda = (\lambda I - A)^{-1}. \tag{12}$$

One can develop a theory for such operator valued functions of a complex variable which parallels very closely the classical complex variable theory. In particular, the concepts of continuity, analyticity and contour integral can be defined for such functions. In many cases the proof of a theorem from classical complex variable theory can be adopted to the operator valued case simply by reinterpreting the meaning of the symbols and systematically replacing the symbol for absolute value by the symbol for the norm of the operator. It is known that R is analytic as an operator valued function.

Let f be a function which is defined and analytic, in the classical complex variable sense on an open set containing $\sigma(T)$. The operational calculus associates an operator $f(A)$ in $[X]$ with f by

$$f(A) = \frac{1}{2\pi i}\int_r f(\lambda)R_\lambda d\lambda, \tag{13}$$

where Γ is the boundary of a Cauchy domain [**12**, p. 288] and Γ lies entirely in $\rho(T)$ and the domain of f and contains $\sigma(T)$ in its interior.

The integration is performed in the positive sense, that is, if one were walking around one of the curves composing Γ the interior of the open set would be on the left. The operator $f(A)$ is independent of the particular curves forming Γ.

There are two standard theorems which will prove useful in our study of the spectrum of L-forms.

THEOREM 3. *Let X be a Banach space and $T \in [X]$. If f and g are two functions defined and analytic on an open set containing $\sigma(A)$, then*

$$(f+g)(A) = f(A) + g(A) \tag{14}$$

and

$$(f \cdot g)(A) = f(A)g(A). \tag{15}$$

Theorem 3 expresses the very useful fact that the mapping $f \to f(T)$ is a ring homomorphism. Equation (14) follows from the linearity of the contour integral in (13). However, some special properties of R_λ are needed to establish (15).

A very interesting fact about the operator $f(A)$ is that its spectrum is uniquely determined by the spectrum of A.

THEOREM 4. *If f is analytic in an open set containing $\sigma(A)$, then*

$$\sigma(f(A)) = f(\sigma(A)), \tag{16}$$

where $f(A) \in [X]$ defined by equation (13) *and $\sigma(f(A))$ denotes the spectrum of $f(A)$ while $f(\sigma(A))$ is the set of complex numbers $f(\lambda)$ with $\lambda \in \sigma(A)$.*

Theorem 4 is often called the spectral mapping theorem; it reduces the problem of finding $\sigma(f(A))$ to a problem in conformal mapping.

The proofs of the above theorem as well as a more detailed discussion of the operational calculus can be found in Taylor [12, 5.6, p. 287].

We shall now apply the operational calculus to study the spectrum of what Toeplitz called regular L-forms.

DEFINITION 5. *An L-form, determined by the sequence $\{a_n\}$, is defined to be regular provided there exists real positive numbers r and R with $r<1<R$ such that the Laurent series*

$$f(\lambda) = \sum_{n=-\infty}^{\infty} a_n \lambda^n \tag{17}$$

converges in the annular region $r < |\lambda| < R$.

From complex variable theory an L-form is regular if and only if the function $f(\lambda)$ in equation (17) is defined and analytic in the annular region containing the unit circle. We will need this fact to obtain the spectrum of a regular L-form.

LEMMA 4. *Let L be a regular L-form and let f be the function associated with L by means of* (17). *Then $L=f(T)$ where T is the L-form in Theorem* 2.2.

Proof. By the operational calculus (cf. equation (13)) we have

$$f(T) = -\frac{1}{2\pi i}\int_{c_1} f(\lambda)R_\lambda d\lambda + \frac{1}{2\pi i}\int_{c_2} f(\lambda)R_\lambda d\lambda, \tag{18}$$

where c_1 is a circle of radius r_1 with $r < r_1 < 1$ and c_2 is a circle of radius R_2 with $1 < R_2 < R$ (both circles about the origin). The appearance of the minus sign in the first integral in equation (18) is due to the fact that we are integrating counterclockwise. In the first integral we have

$$(19)\quad R_\lambda = -[I - \lambda T^{-1}]^{-1} = -(T^{-1} + \lambda T^{-2} + \lambda^2 T^{-3} + \cdots + \lambda^n T^{-(n+1)} + \cdots)$$

and the infinite sum converges in the uniform operator topology since $\|\lambda T^{-1}\| < 1$ for $|\lambda| < 1$. Similarly for the second integral in equation (18) we have

$$(20)\qquad R_\lambda = [\lambda(1 - T/\lambda)]^{-1} = 1/\lambda I + 1/\lambda^2 T + \cdots + 1/\lambda^{n+1} T^n + \cdots$$

and the infinite sum appearing on the right also converges since $\|T/\lambda\| < 1$. The reader can easily verify that

$$(21)\qquad T^{-n} = -\frac{1}{2\pi i}\int_{c_1} \lambda^{-n} R_\lambda d\lambda \qquad (n = 1, 2, 3, \cdots),$$

and that

$$(22)\qquad T^n = \frac{1}{2\pi i}\int_{c_2} \lambda^n R_\lambda d\lambda \qquad (n = 0, 1, 2, \cdots).$$

Integrating in equation (18) term by term it follows by the residue theorem that $f(T) = \sum_{-\infty}^{\infty} a_n T^n$.

The proof of Lemma 4 can be completed by using Lemma 5.

LEMMA 5. *Let L be an L-form and let T be the L-form in Theorem 2. Then (even if L is not a regular L-form) the series $\sum_{-\infty}^{\infty} a_k T^k$ converges in the uniform operator topology on $[l_p]$ for $1 \leqq p \leqq \infty$ and*

$$(23)\qquad L = \sum_{-\infty}^{\infty} a_k T^k.$$

Proof. Let $S_n = \sum_{k=0}^{n} a_k T^k$; then

$$\|S_n - S_m\| \leqq \sum_{k=m+1}^{n} |a_k| \, \|T^k\| = \sum_{k=m+1}^{n} |a_k|,$$

where we have used the fact that $\|T^k\| = 1$. Since $\sum |a_n| < \infty$ we see that the sequence $\{S_n\}$ is a Cauchy sequence in $[l_p]$. Since $[l_p]$ is complete (cf. Lemma 1) the sequence $\{S_n\}$ is convergent. Hence the series $\sum_{k=0}^{\infty} a_k T^k$ is convergent in the uniform operator topology. A similar argument shows that the series of negative terms converges.

Equation (23) can be established by examining the matrix for the series on the right hand side.

COROLLARY.

$$(24)\qquad \|L\| \leqq \sum_{k=-\infty}^{\infty} |a_k|.$$

Note that we refrain from writing $L=f(T)$ if L is not known to be regular; we reserve this notation for elements of the operational calculus.

THEOREM 5. *Let L be a regular L-form. Then $L\in[l_p]$ and the spectrum of L is*

$$\sigma(L) = \{f(\lambda): |\lambda| = 1\}, \tag{25}$$

where $f(\lambda)$ is defined by equation (17).

Proof. Let T be the L-form described in Theorem 2 with $\rho(T)=\{\lambda: |\lambda|=1\}$. Since the function f is analytic on a neighborhood of $\sigma(T)$, the operational calculus defines a function $f(T)$ in $[l_p]$. However, $L=f(T)$ by Lemma 5, hence the spectral mapping theorem (Theorem 4) yields the assertion about $\sigma(L)$ and shows $L\in[l_p]$.

Theorem 5 for the case $p=2$ is due to Toeplitz [9]. For the cases $1<p<\infty$, Theorem 5 is a corollary of a theorem of Krabbe [7, p. 42]. See also [4].

By considering the effect of L on a vector x for which $x_0=1$ and $x_i=0$ if $i\neq 0$ it follows from the definition of the norm that

$$\left(\sum_{k=-\infty}^{\infty} |a_k|^p\right)^{1/p} \leqq \|L\|_p, \tag{26}$$

where we have written $\|L\|_p$ for the norm of L as an operator of l_p. In (26) we need only consider values of p such that $1\leqq p\leqq 2$ since the norm of the transpose of L as an operator on the dual space is the same as the norm of L. From Theorem 5 (or Theorem 11 if L is not regular) and the fact that the spectral radius, $\sup|\lambda|$ for $\lambda\in\sigma(L)$, is not greater than the norm we see that

$$\sup_{|\lambda|=1} \left|\sum a_n\lambda^n\right| \leqq \|L\|_p. \tag{27}$$

From (26) and (24) we deduce that $\|L\|_1=\|L\|_\infty=\sum|a_n|$ and from (24) and (27) that $\|L\|_p=\sum|a_n|$ if all the $a_n\geqq 0$ (or if all the a_n are rotated by multiplying by $e^{i\theta}$.)

There are two more consequences of Lemma 4 which will be useful later.

THEOREM 6. *Let L_f and L_g denote any two L-forms corresponding to the functions f and g respectively, according to equation* (17). *Then*

$$L_fL_g = L_gL_f = L_{fg}. \tag{28}$$

Proof. We begin by showing the theorem is true for regular L-forms L_f and L_g. Since $L_f=f(T)$ and $L_g=g(T)$ by Lemma 4, the operational calculus assures us that $f(T)g(T)=g(T)f(T)$ since $f(\lambda)g(\lambda)=g(\lambda)f(\lambda)$ and so the first equality in (28) is valid. Also, by equation (15), we have $(f\circ g)(T)=f(T)g(T)$ or $L_{fg}=L_fL_g$.

The theorem can now be proved for arbitrary L-forms by using the fact that every L-form is a limit of regular L-forms, the partial sums of the series defining f, in the uniform operator topology of $[l_p]$ by Lemma 5.

Theorem 6 could be established by performing the matrix multiplication but this procedure is somewhat tedious.

We intend to extend some of the above results concerning the spectrum of regular L-forms to L-forms which are not necessarily regular. We postpone this discussion, however, until Section 5 since the situation is a little more complicated than it might first appear (cf. first paragraph in Section 5).

4. The fine structure spectral mapping theorem. In this section we plan to study what is called the fine structure of the spectrum of a regular L-form.

If λ is an eigenvalue of an L-form L, then λ is in the spectrum of L for, just as in the finite dimensional case, $L-\lambda I$ is not 1:1, i.e., there exists no inverse (cf. Theorem 2). However, for infinite dimensional spaces $L-\lambda I$ can be 1:1 and it is still possible to have λ in $\sigma(L)$. For by the very definition of an in finite set the range of $L-\lambda I$ need not be the whole space and so $(L-\lambda I)^{-1}$ is not an element of $[X]$. Thus there are a number of ways a complex number λ can get into the spectrum of an operator. Points in the spectrum are classified according to these various possibilities. This study is known as fine structure theory.

The classification scheme, we wish to introduce, is based on the idea that there are basically three ways that a linear operator A in $[X]$ may fail to have an inverse A^{-1} in $[X]$.

DEFINITION 6. *Let X be a Banach space and suppose $A\in[X]$. Then the properties P_i ($i=1, 2, 3$) are defined as follows:*

P_1: *A is not* 1:1.

P_2: *The range of A is not dense in X.*

P_3: *There exists a sequence of unit vectors $x_n(\|x_n\|=1)$ for which $\|Ax_n\|\to 0$ as $n\to\infty$.*

P_1 states that A^{-1} does not exist; the meaning of P_2 is apparent. The property P_3 means that A does not have a continuous inverse (cf. Taylor Theorem 3.1-B [**12**, p. 56]).

From the above interpretations of the properties P_i it is evident that a point $\lambda\in\sigma(L)$ if and only if $L-\lambda I$ has at least one of the properties P_i. To illustrate that these various possibilities actually occur we study the fine structure of the operator T.

THEOREM 7. *If $T: l_p\to l_p$ is the L-form of Theorem 2, then the fine structure of the spectrum of T ($\sigma(T)=\{\lambda: |\lambda|=1\}$) is as follows:*

(a) *$p=\infty$, $(\lambda I-T)$ has the properties P_1 and P_2;*

(b) *$1<p<\infty$. $(\lambda I-T)$ has the property P_3 but neither P_2 nor P_1;*

(c) *$p=1$. $(\lambda I-T)$ has the properties P_2 and P_3.*

Proof. (a) We first prove that the range of $\lambda I-T$ is not dense in l_∞. Let $\omega=(\cdots, \lambda^{-n}, \cdots, \lambda^{-1}, 1, \lambda, \cdots, \lambda^n, \cdots)$; then $\omega\in l_\infty$ and if $\|y-\omega\|=\epsilon<1$, then y is not in the range of $\lambda-T$. Assume the contrary and suppose that $y=(\lambda-T)x$; set $y=\omega+\theta$, where $\theta=\{\theta_n\}$ and $|\theta_n|<\epsilon$ for an arbitrary $\epsilon>0$. Then from equation (8) we obtain

$$x_k = \lambda^k x_0 - (\lambda^{k-1}\theta_0 + \lambda^{k-2}\theta_1 + \cdots + \lambda\theta_{k-2} + \theta_{k-1}) - k\lambda^{k-1},$$

which implies that $|k\lambda^{k-1}+x_k-\lambda^k x_0| \leqq k\epsilon$. Upon dividing by k in the last inequality and letting $k\to\infty$ we obtain $1\leqq\epsilon$. A similar contradiction arises upon using (9).

The inverse does not exist since we have an eigenvalue $\{\lambda^k x_0\}$ as we have seen in Theorem 2.

(b) The results in (a) show that an eigenvector must be of the form $\{\lambda^k x_0\}$. However, since $|\lambda|=1$, this element does not belong to $l_p(1\leqq p<\infty)$ unless $x_0=0$ and so $x=0$. Since $(\lambda I-T)x=0$ implies $x=0$, the inverse exists.

Consider the vectors

$$x_n = (\cdots 0, 0, \cdots, \lambda^{-n}, \cdots \lambda^{-1}, 1, \lambda, \lambda^2, \cdots \lambda^n, 0, 0 \cdots);$$

then if $(\lambda I-T)x_n=y_n$, all the components of y_n are zero except the nth and $-(n+1)$th which are λ^n and $-\lambda^{-n}$ respectively.

From this it follows that $\|(\lambda-T)x_n\|\leqq 2$ for all n while $\|x_n\|\to\infty$ which yields the assertion that $(\lambda I-T)^{-1}$ is not continuous.

Finally we will prove that the range of $\lambda I-T$ is dense in l_p. In the proof of Theorem 2 we saw that the range of $\lambda I-T$ was contained in the hyperplane given by equation (10). We will show that each sequence with only a finite number of nonzero terms is a limit of elements satisfying equation (11), hence equation (10) since $|\lambda|=1$. Consider the basis vector u with $u_0=1$ and $u_k=0$ $(k\neq 0)$. It is the limit of the sequence $\{\xi_k^{(n)}\}$ where $\xi_0^{(n)}=1-1/n$, $\xi_k^{(n)}=-1/n$ $(k=1, 2, \cdots, n-1)$, and $\xi_k^{(n)}=0$ otherwise. For $\|u-\xi^{(n)}\|=(n|1/n|^p)^{1/p}$ which approaches zero as $n\to\infty$ if $p>1$ as the reader can easily verify.

(c) The argument used in (b) shows that the inverse exists and is not continuous, but the solutions of (11) are not dense in l_1. We will show that no sequence $\{\xi^{(n)}\}$ satisfying (11) approaches $u_0=1$ and $u_k=0$ $(k\neq 0)$. Assume that $\xi^{(n)}\to u$ and, without loss of generality, we may suppose that $\|\xi^{(n)}\|=1$. Since

$$\sum_{-\infty}^{\infty} \xi_k^{(n)} = 0, \qquad \sum_{k\neq 0} \xi_k^{(n)} \to -1$$

as $n\to\infty$ since $\xi_0^{(n)}\to 1$. Because $\xi_0^{(n)}\to 1$ and $\|\xi^{(n)}\|=1$, $\sum_{k=0}\xi_k^{(n)}\to 0$ and we have a contradiction.

There is an interesting aspect to the proof of the above theorem. From the fact that no hyperplane is dense in Euclidean n-space most of us would guess that no hyperplane is dense in an infinite dimensional space. Part (b) shows, however, that indeed the points of $l_p(1<p\leqq\infty)$ lying on the hyperplane (11) are dense in l_p.

We are now in a position to study the fine structure of the spectrum of an operator obtained by means of the operational calculus.

THEOREM 8. (The fine structure spectral mapping theorem.) *Let X be a Banach space, suppose $A\in[X]$, and let f be analytic on an open set containing the spectrum of A. If $\alpha I-A$ has the property P_i, for $i=1, 2, 3$, then so has $f(\alpha)I-f(A)$. If $\mu I-f(A)$ has the property P_i, for $i=1, 2, 3$, and if $f(\alpha)\neq\mu$ on each component*

of the domain of f, then there is an α in $\sigma(A)$ such that $f(\alpha)=\mu$ and $\alpha I-A$ has the property P_i.

This theorem is proved in Hille and Phillips [5, pp. 204–5]. Essentially Theorem 8 says that every "bad" property of $\alpha I-A$ is passed on to $f(\alpha)I-f(A)$. Conversely for nonconstant f, every "bad" property of $\alpha I-f(A)$ must have been inherited from $\alpha I-A$ for some α such that $f(\alpha)=\mu$. It is essentially this genetic-like character of the properties P_i which makes the operational calculus so useful.

The above theorem will be applied to find the spectrum of a regular L-form. We first take care of the exceptional case—namely, $L=\alpha I$ where α is an arbitrary complex number. Here the spectrum of L is the point $\lambda=\alpha$ and $\lambda I-L$ has the properties P_1, P_2, and P_3 at this point.

THEOREM 9. *Let $L\colon l_p\to l_p$ be a regular L-form (Def. 5) which is not a multiple of the identity then the fine structure of the spectrum of L ($\sigma(L)=\{f(\lambda)\colon |\lambda|=1\}$) is as follows:*

(a) $p=\infty$, $(\lambda I-L)$ *has the properties P_1 and P_2.*

(b) $1<p<\infty$, $(\lambda I-L)$ *has the property P_3 but neither P_1 nor P_2.*

(c) $p=1$, $(\lambda I-L)$ *has the properties P_2 and P_3.*

The proof is an immediate consequence of the spectral mapping theorem, Theorem 8, and the properties of the spectrum of T given in Theorem 7, since $L=f(T)$ by Lemma 4. In Theorem 9, the assertions about the property P_1 were proved by Krabbe [8, p. 784] for $1<p<\infty$.

5. Newburgh's theorem. We would now like to study the spectrum of the general L-form. One might expect that this would be directly obtainable from the fact that the general L-form is a limit of regular L-forms. Unfortunately, that $T_n\to T$ in norm does not necessarily imply $\sigma(T_n)\to\sigma(T)$. For an example of this situation see Rickart [11, p. 282].

Newburgh [10, Theorem 4, page 168] proved the following result:

THEOREM 10. *If in a Banach algebra $T_i\to T$ and $T_iT=TT_i$ for all i, then* $\lim \sigma(T_i)=\sigma(T)$.

The next theorem was obtained by Krabbe [9] using Wiener's theorem. We give our own proof in order to obtain Wiener's theorem as a corollary.

THEOREM 11. *Let L be an L-form determined by the sequence $\{a_k\}$. Then the spectrum of L as an operator from l_p into l_p is*

$$\sigma(L) = \{f(\lambda)\colon |\lambda| = 1\}, \tag{29}$$

where f is defined by $f(\lambda)=\sum_{-\infty}^{\infty} a_k\lambda^k$ for $|\lambda|=1$.

Proof. Let $L_n=f_n(A)$ where $f_n(\lambda)=\sum_{k=-n}^{n} a_k\lambda^k$; then $L_n\to L$ by Lemma 5. Theorem 6 assures us that $L_nL=LL_n$ for all n. Since $[l_p]$ is a Banach algebra,

Theorem 10 applies and the result follows from Theorem 5, equation (25) since $f_n(\lambda)\to f(\lambda)$ for each λ on the unit circle.

Since the series defining f is uniformly convergent on the unit circle (because of our hypothesis that $\sum|a_k|$ is convergent) the function f is continuous. Thus we have the following corollary to Theorem 11.

THEOREM 12. *Let L be an L-form. Then the spectrum of L is connected.*

The form of the spectrum of L and some well-known theorems [**12**, p. 281, Theorem 5.5 F-G] enable us to deduce the next result.

THEOREM 13. *An L-form is completely continuous if, and only if, it is the zero operator.*

In Section 4 we proved that $L-\lambda I$ cannot have a continuous inverse if $\lambda\in\sigma(L)$, that is, $L-\lambda I$ has the property P_3 for all $\lambda\in\sigma(L)$. This result holds for any L-form L and can be proved by using the fact that L is a limit of regular L-forms together with know theorems from any one of the following three references: Kato [**6**], Gohberg and Krein [**3**], or Gindler and Taylor [**2**].

We have not been able to extend to the general case, the remaining results in Section 4 concerning the fine structure of the spectrum of a regular L-form. For an interesting discussion of some of the problems involved in determining the fine structure of certain L-forms see [**4**, p. 48].

6. The famous theorem of Wiener. A (to us) startling and gratifying application of the study of the spectrum of L-forms is a new proof of Wiener's theorem [**14**]. Anyone who thinks this theorem is obvious should examine Wiener's original proof. Gelfand [**1**] proved this theorem using the Gelfand theory for Banach algebras. Our proof depends on this theory through our use of Newburgh's result (Theorem 10) in obtaining the spectrum of an L-form. Our proof has the advantage that we do not need to study the structure of the maximal ideals in the space of absolutely convergent Fourier series (cf. [**5**, p. 136]).

THEOREM 14. *Let f be a function whose Fourier series $\sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ is absolutely convergent. Then, if f is never zero on the unit circle, the reciprocal of f has an absolutely convergent Fourier series.*

Proof. Let L be the L-form corresponding to f. Then $0\notin\sigma(L)$ by Theorem 2 and so L has an inverse $L^{-1}\in[l_1]$. The theorem will be proved if we succeed in showing that L^{-1} is the L-form corresponding to $1/f$.

To see that L^{-1} is indeed the L-form of some function g, let g be the function whose Fourier series has coefficients $L^{-1}u$, where $u_0=1$ and $u_n=0$ otherwise. (Observe that this is just the "central" column in the L-form.) Then $g(\lambda) = \sum b_n\lambda^n$ and we must show $\sum|b_n|<\infty$ which is equivalent to the fact that g has an absolutely convergent Fourier series. However, this result is a consequence of $\sum|b_n| = \|L^{-1}\|$.

It remains to show that $g=1/f$. Now $L_fL_g=L_{fg}$ and any L-form is completely determined by one column in the matrix. Hence, $L_{fg}u=L_fL_gu=LL^{-1}u=u$ implies $fg=1$.

We leave it for the reader to verify that the following theorems on absolutely convergent Fourier series are also corollaries of the theory of L-forms. For more classical proofs of these theorems see Zygmund [**15**, pp. 245–7].

THEOREM 15. *Let f be a function with absolutely convergent Fourier series and suppose that ϕ is a function which is analytic on a neighborhood of the set* $\{f(\lambda): |\lambda|=1\}$. *Then the composition $\phi\circ f$ has an absolutely convergent Fourier series.*

THEOREM 16. *Let f be a function with absolutely convergent Fourier series and let F_n be a sequence of analytic functions converging uniformly to* 0 *in a neighborhood of the range of f. Then $\|F_n\circ f\|\to 0$ where $\|g\|=\sum|a_n|$ if the a_n are the Fourier coefficients of g.*

THEOREM 17. *If f has an absolutely convergent Fourier series, then* $\lim_{n\to\infty}\|f^n\|^{1/n}=\sup_{|x|=1}|f(x)|$.

The proof of Theorem 17 by means of L-forms requires a theorem concerning the spectral radius [**12**, Theorem 5.2-E, p. 263].

Acknowledgments. The authors wish to acknowledge their indebtedness to Mr. Neal Koblitz. We are also indebted to the referee for several helpful suggestions.

References

1. I. M. Gelfand, Über absolut konvergente trigonometrische Reihen und Integrale, Mat. Sbornik, 9 (1941) 51–66.

2. H. A. Gindler and A. E. Taylor, The minimum modulus of a linear operator and its use in spectral theory, Studia Mathematica, 22 (1962) 15–41.

3. I. C. Gohberg and M. G. Krein, The basic propositions on defect numbers, root numbers and indices of linear operators, Amer. Math. Soc. Translations, Series 2, 13 (1960) 185–264.

4. P. Hartman, On Laurent operators on l_p, Proc. Amer. Math. Soc., 8 (1957) 45–48.

5. E. Hille and R. S. Phillips, Functional Analysis and Semi-groups, Amer. Math. Soc., Colloq. Publ., 31 (revised ed.) 1957.

6. T. Kato, Perturbation theory for nullity, deficiency and other quantities of linear operators, J. Analyse Math., VI (1958) 261–322.

7. G. L. Krabbe, Spectral isomorphisms for some rings of matrices on a Banach space, Amer. J. Math., 78 (1956) 42–50.

8. ———, Abelian rings and spectra of operators on l_p, Proc. Amer. Math. Soc., 7 (1956) 783–790.

9. ———, On the spectra of certain Laurent matrices, Proc. Amer. Math. Soc., 8 (1957) 894–897.

10. J. D. Newburgh, The variation of spectra, Duke Math. J., 18 (1951) 165–176.

11. C. E. Rickart, General Theory of Banach Algebras, Van Nostrand, Princeton, 1960.

12. A. E. Taylor, Introduction to Functional Analysis, Wiley, New York, 1958.

13. O. Toeplitz, Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen veränderlichen. I. Teil: Theorie der L-Formen., Math. Annalen, 70 (1911) 351–376.

14. N. Wiener, Tauberian theorems, Annals of Math., (2) 33 (1932) 1–100.

15. A. Zygmund, Trigonometric Series I, University Press, Cambridge, 1959.

ALGEBRAIC PROPERTIES OF CERTAIN INTEGRAL TRANSFORMS

RAYMOND REDHEFFER, University of California, Los Angeles

1. Introduction. Many important investigations lead to integrals of the type

$$T_a f = x^{a-1} \int_0^x s^{-a} f(s)\, ds, \qquad \tilde{T}_a f = x^{a-1} \int_x^\infty s^{-a} f(s)\, ds, \tag{1}$$

where a is constant. In the theory of entire functions, these integrals are needed for estimation of canonical products. They also occur in the representation of special functions; for instance,

$$x^{a-1}\gamma(1 - a, x) = T_a(e^{-s}),$$

where γ is the incomplete gamma function. The substitution $s = e^t$ gives expressions related to the Laplace transform; e.g.,

$$x^{1-a}(T_a f + \tilde{T}_a f) = \int_{-\infty}^{\infty} e^{-ta} f(e^t) e^t\, dt.$$

Our purpose here is to indicate some curious algebraic properties of the transforms (1). In the course of the discussion we suggest a number of research problems, suitable for an alert class in advanced calculus.

2. The product formulas. To avoid convergence difficulties let f be continuous and let $f(s) = 0$ near $s = 0$. Then a change in the order of integration gives

$$\begin{aligned} T_a(T_b f) &= x^{a-1} \int_0^x t^{-a} \left[t^{b-1} \int_0^t s^{-b} f(s)\, ds \right] dt \\ &= x^{a-1} \int_0^x \frac{x^{b-a} - s^{b-a}}{b - a}\, s^{-b} f(s)\, ds. \end{aligned}$$

If we define sums, products and scalar multiples of transformations in a natural manner, the foregoing relation reduces to

$$T_a T_b = \frac{T_a - T_b}{a - b}, \qquad (a \neq b). \tag{2}$$

We leave it to the reader to show that

$$\tilde{T}_a \tilde{T}_b = \frac{\tilde{T}_a - \tilde{T}_b}{b - a}, \qquad (a \neq b), \tag{3}$$

if f vanishes near ∞, and that

$$T_a \tilde{T}_b = \tilde{T}_b T_a = \frac{T_a + \tilde{T}_b}{b - a}, \qquad (b > a), \tag{4}$$

if f vanishes both near 0 and near ∞.

Since our transformations are linear, the distributive law

$$T_a(T_b + T_c) = T_aT_b + T_aT_c$$

holds, and extended products can be computed by repeated application of the foregoing relations. In particular, $(T_aT_b)T_c$ is

$$\frac{T_a}{(a-b)(a-c)} + \frac{T_b}{(b-a)(b-c)} + \frac{T_c}{(c-a)(c-b)}$$

for a, b, c all unequal. The same result is obtained for $T_a(T_bT_c)$ (and that is just as well, because multiplication of transformations is always associative).

3. An algebra without squares. The formula (2) shows two surprising facts: the product is representable in terms of the sum, and it is commutative. Since the formula breaks down when $b=a$ we have no quarrel with those who say, "You can't form x^2y, because x was already used up in forming xy." In the present discussion, repeated factors are really excluded.

To develop an algebra subject to this limitation, start with a set of real numbers M, and form the various transformations T_a, with $a \in M$. These give a basis. The elements themselves are linear combinations, that is, expressions of form $\sum \alpha(a)T_a$, where α is a real-valued function defined on M, and a runs over M. Different functions α give different elements. The sum of two elements is defined in an obvious way, and by (2) the product is

$$[\sum \alpha(a)T_a][\sum \beta(b)T_b] = \sum\sum \frac{\alpha(a)\beta(b)}{a-b}(T_a - T_b). \tag{5}$$

This requires that $\alpha(a)\beta(b)=0$ for $a=b$; in other words, the sets $\{T_a\}$ and $\{T_b\}$ must be disjoint. Then addition is commutative and associative, multiplication is commutative and associative, and the distributive law holds. But we cannot form the square of a basis element T_a, or any product which leads to such a square.

4. Characteristic functions. Upon letting $b \to a$ in (2) one is led to conjecture that

$$T_a^2 = \frac{\partial}{\partial a} T_a,$$

or, more explicitly,

$$T_a^2 f = \frac{\partial}{\partial a} T_a f = (\log x)T_a[f(s)] - T_a[f(s)\log s].$$

The latter relation can be verified directly as in the derivation of (2). But a more efficient method is to note that x^m is a characteristic function for T_a, whenever the constant $m > a-1$. That is,

$$T_a x^m = \lambda x^m, \qquad \lambda = (m - a + 1)^{-1}. \tag{6}$$

This shows that $T_a^n x^m = \lambda^n x^m$, and hence the formula

$$T_a^{n+1} = \frac{1}{n!}\left(\frac{\partial}{\partial a}\right)^n T_a \tag{7}$$

holds when both sides are applied to $f(x) = x^m$. By linearity the formula must then hold for all polynomials f of degree greater than $a-1$ in the lowest terms. Since a continuous function vanishing near 0 can be approximated by such polynomials, (7) is established for the whole class $\{f\}$ with which we are concerned. It is left for the reader to re-establish (5) in the same way, and to evaluate more general products, such as $T_a^3 T_b^2$.

Equation (7) leads to a variety of symbolic formulas, of which

$$T_a^{-1} e^D T_a = (1 - T_a)^{-1}, \qquad D = \frac{\partial}{\partial a}$$

is a typical example. Construction of other such formulas and investigation of their validity is left to the reader.

5. Multipliers preserving positivity. We define a generalized polynomial by

$$f(x) = \sum_{m \in M} c(m) x^m,$$

where m ranges over a given finite set, M, of nonnegative real numbers. The notation emphasizes that the exponents need not be integers. A function ϕ is called a *multiplier* if the condition $f(x) \geqq 0$ for $0 \leqq x \leqq x_0$ implies

$$\sum_{m \in M} \phi(m) c(m) x^m \geqq 0, \qquad 0 \leqq x \leqq x_0.$$

Choosing $y = a - 1$ in (6) we find that $\phi(m) = (m-y)^{-n}$ is a multiplier, where n is a positive integer and y is any constant smaller than all the m's. Superposition over y leads to additional multipliers, of form

$$\phi(m) = \int_{-\infty}^{y_0} (m - y)^{-n} \, d\mu(y), \qquad d\mu \geqq 0. \tag{8}$$

Without reference to T_a it is evident that $\phi(m) = y^m$ is also a multiplier, provided $0 \leqq y \leqq 1$. This leads to the class

$$\phi(m) = \int_0^1 y^m \, d\mu(y), \qquad d\mu \geqq 0.$$

We leave it to the reader to investigate the relation between these two classes and to decide whether the condition $n =$ integer in (8) is really necessary.

6. Convergence. Let f be a continuous nonnegative function vanishing near 0 and define $F = T_a f$, where a is any constant. We shall prove

$$(9)\qquad \int_0^\infty \frac{F(x)}{x^{a+1}}\,dx = \int_0^\infty \frac{f(x)}{x^{a+1}}\,dx$$

in the sense that if either integral converges, the other does too, and the values are equal.

Integration by parts gives

$$\int_0^R \frac{F(x)}{x^{a+1}}\,dx = \int_0^R \frac{f(x)}{x^{a+1}}\,dx - E,$$

where the error E is

$$E = \frac{1}{R}\int_0^R \frac{f(x)}{x^a}\,dx.$$

Let us write (f conv) as an abbreviation for the statement: "The integral involving f is convergent as $R \to \infty$," and similarly for F. Then (9) is established if it can be shown that

$$(f \text{ conv}) \Rightarrow (E \to 0) \Rightarrow (F \text{ conv}) \Rightarrow (f \text{ conv}).$$

But the first implication follows from

$$E = \int_0^R \left(\frac{x}{R}\right)\frac{f(x)}{x^{a+1}}\,dx \leqq \delta \int_0^{\delta R} \frac{f(x)}{x^{a+1}}\,dx + \int_{\delta R}^R \frac{f(x)}{x^{a+1}}\,dx,$$

where δ is a small positive constant. The second is evident, and the third follows from

$$\int_0^R \frac{f(x)}{x^{a+1}}\left(1 - \frac{x}{R}\right)dx \geqq \int_0^{R/2} \frac{f(x)}{x^{a+1}}\left(\frac{1}{2}\right)dx.$$

Under the same hypothesis on f let $\tilde{F} = \tilde{T}_a f$. Then by a similar argument

$$(10)\qquad \int_0^\infty \frac{\tilde{F}(x)}{x^{a-1}}\,dx = \int_0^\infty \frac{f(x)}{x^{a-1}}\,dx,$$

or both integrals are divergent, and

$$(11)\qquad \int_0^\infty \frac{\tilde{F}(x)}{x^a}\,dx = \int_0^\infty \frac{f(x)}{x^a}\log x\,dx,$$

or both are divergent.

Equations (2)–(4) and the distributive law give

$$(12)\qquad TT_a = \frac{T}{b-a}\quad (b > a), \qquad T\tilde{T}_a = \frac{T}{a-b}\quad (a > b),$$

where $T=T_b+\tilde{T}_b$. The special case $b=a+1$ in the first relation (12) gives (9), and the special case $b=a-1$ in the second gives (10). We leave it to the reader to analyze (3)–(4) after the manner of (9), and to fit (11) into the same context. What happens if the behavior of f near 0 (and not only near ∞) is left unspecified? How far can one relax the conditions of positivity and continuity?

7. Divergence. Unless f is severely restricted, the condition $b>a$ in

$$T_a\tilde{T}_b = \tilde{T}_bT_a = \frac{T_a+\tilde{T}_b}{b-a} \tag{13}$$

is essential to ensure convergence. On the other hand, for broad classes of functions f the right side of (13) is an analytic function of a, provided $a\neq b$. Thus the right side gives the analytic continuation of the left side. By analogy to Euler's method of summing divergent series, we therefore define the left side by the right side even when $b<a$. (If $f>0$, the resulting value is always negative. Perhaps it is unfair to define an integral as -1, say, when the integral diverges to $+\infty$; but the series $1+2+4+8+\cdots$ has long been subjected to the same indignity.)

Dropping the condition $b>a$ in (4), we find that our multiplication table takes the form

	T_a	$\tilde{T}_a$
T_b	T_a-T_b	$\tilde{T}_a+T_b$
$\tilde{T}_b$	$-T_a-\tilde{T}_b$	$-\tilde{T}_a+\tilde{T}_b$

with the understanding that each entry must be divided by $a-b$. It is possible to generate all the entries from any one of them by the following rule: Whenever we add or drop the $\sim$ on one term, we change the sign of the other term. The reader is invited to find other algebras with the same multiplication table, and to interpret the formulas

$$T_a^2 \doteq \tilde{T}_a^2 \doteq 0, \qquad T_a\tilde{T}_a \doteq -\tilde{T}_aT_a \doteq T_a+\tilde{T}_a$$

that are obtained by ignoring the factor $a-b$.

The preparation of this paper was sponsored by the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

A NOTE ON SELFADJOINT OPERATORS

L. CARLITZ, Duke University

1. A. M. Krall [**1**, **2**] has proved that the operator

$$L = AD^n + \sum_{j=0}^{n-1} \binom{n}{j} \overline{C}_{n-j} A^{(n-j)} D^j \qquad \left(D = \frac{d}{dx}\right)$$

satisfies the equation

$$L = (-1)^n M, \tag{1}$$

where $M = -(1)^n D^n A + \sum_{j=0}^{n-1} (-1)^j \binom{n}{j} \overline{C}_{n-j} D^j A^{(n-j)}$. Here

$$\overline{C}_{r-1} = 2(2^r - 1) \frac{B_r}{r} \qquad (r = 2, 3, 4, \cdots)$$

and the B_r are the Bernoulli numbers in the even suffix notation.

The presence of the Bernoulli numbers in these formulas is a bit surprising. The situation becomes somewhat clearer in Nörlund's notation [**3**, Ch. 2]. Put

$$C_{n-1} = 2^n(1 - 2^n) \frac{B_n}{n} \qquad (n = 1, 2, 3, \cdots),$$

so that $C_n = -2^n \overline{C}_n$, $(n = 1, 2, 3, \cdots)$. Now define the differential operators δ_a, δ_y by means of

$$\delta_a A = A', \qquad \delta_a y = y\delta_a, \qquad \delta_y A = A\delta_y, \qquad \delta_y y = y',$$

where y is an arbitrary function of x. Then

$$D^j = \sum_{s=0}^{j} \binom{j}{s} \delta_a^{j-s} \delta_y^s = (\delta_a + \delta_y)^j$$

and

$$L = [2\delta_y^n - (\tfrac{1}{2}C\delta_a + \delta_y)^n]A,$$
$$(-1)^n M = [2(\delta_a + \delta_y)^n - (-\tfrac{1}{2}C\delta_a + \delta_a + \delta_y)^n]A.$$

Since $C_{2r} = 0$, $(r = 1, 2, 3, \cdots)$, the latter formula may be replaced by $(-1)^n M = (\frac{1}{2}C\delta_a + \delta_a + \delta_y)^n A$.

It follows that (1) is equivalent to

$$2v^n = (\tfrac{1}{2}Cu + v)^n + (\tfrac{1}{2}Cu + u + v)^n, \tag{2}$$

where u, v are arbitrary symbols such that $uv = vu$. In view of [**3**, p. 28]

$$(C + 2)^n + C^n = \begin{cases} 2 & (n = 0) \\ 0 & (n > 0), \end{cases}$$

(2) is immediate.

2. To get a result like (2) involving the B_n we recall that $(B+1)^n = B^n$, $(n \neq 1)$, which implies

$$(Bu + u + v)^n = (Bu + v)^n + nuv^{n-1}.$$

Hence

$$(3) \qquad (B\delta_a + \delta_a + \delta_y)^n = (B\delta_a + \delta_y)^n + n\delta_a\delta_y^{n-1}.$$

If we prefer, (3) may be replaced by the following more explicit operational identity:

$$(4) \qquad \sum_{j=0}^{n} \binom{n}{j} B_{n-j} D^j A^{(n-j)} = \sum_{j=0}^{n} \binom{n}{j} B_{n-j} A^{(n-j)} D^j + nA' D^{n-1}.$$

We may also mention a result containing the Euler numbers E_n which satisfy

$$(5) \qquad (E + 2)^n + E^n = 2.$$

It follows from (6) that $(\frac{1}{2}Eu+u+v)^n + (\frac{1}{2}Eu+v)^n = 2(u+v)^n$, so that

$$(6) \qquad (\tfrac{1}{2}E\delta_a + \delta_a + \delta_y)^n + (\tfrac{1}{2}E\delta_a + \delta_y)^n = 2(\delta_a + \delta_y)^n.$$

Explicitly we have

$$(7) \qquad \sum_{j=0}^{n} \binom{n}{j} 2^{-n+j} E_{n-j} D^j A^{(n-j)} + \sum_{j=0}^{n} \binom{n}{j} 2^{-n+j} E_{n-j} A^{n-j} = 2D^n A.$$

3. More general results can be stated in terms of the Bernoulli and Euler polynomials $B_n(z) = (B+z)^n$, $E_n(z) = (\frac{1}{2}C+z)^n$. We recall that

$$(8) \qquad B_n(z + 1) - B_n(z) = nz^{n-1},$$

$$(9) \qquad E_n(z + 1) + E_n(z) = 2z^n.$$

It follows from (8) that $((B+z)u+u+v)^n = ((B+z)u+v)^n + nu(zu+v)^{n-1}$. Taking $u = \delta_a$, $v = \delta_y$ we get the operational identity

$$(10) \qquad \sum_{j=0}^{n} \binom{n}{j} B_{n-j}(z) D^j A^{(n-j)} = \sum_{j=0}^{n} \binom{n}{j} B_{n-j}(z) A^{(n-j)} D^j + n \sum_{j=0}^{n-1} \binom{n-1}{j} z^{n-j-1} A^{(n-j)} D^j.$$

Similarly it follows from (9) that

$$[(\tfrac{1}{2}C + z)u + v]^n + [(\tfrac{1}{2}C + 2)u + u + v]^n = 2(zu + v)^n,$$

which implies

$$\sum_{j=0}^{n}\binom{n}{j}E_{n-j}(z)A^{(n-j)}D^{j}+\sum_{j=0}^{n}\binom{n}{j}E_{n-j}(z)D^{j}A^{(n-j)} = 2\sum_{j=0}^{n}\binom{n}{j}z^{n-j}A^{(n-j)}D^{j}. \tag{11}$$

Supported in part by NSF grant GP-1593.

References

1. A. M. Krall, Selfadjoint differential expressions, this MONTHLY, 67 (1960) 876–878.
2. ———, Selfadjoint differential expressions of odd order, this MONTHLY, this issue, page 61.
3. N. E. Nörlund, Vorlesungen über Differenzenrechnung, Berlin, 1924, repr. Stechert-Hafner, New York, 1954.

PARTICULAR INTEGRALS FOR LINEAR DIFFERENTIAL EQUATIONS

J. C. BURNS, Australian National University

1. The author recently needed particular integrals for the following pair of linear differential equations:

$$y'' + y = \sin(x + \epsilon), \tag{1}$$

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\nu^2}{x^2}\right)y = J_\nu(x). \tag{2}$$

These are easily found: for (1), the integral $-\frac{1}{2}x\cos(x+\epsilon)$ is well known and standard methods give $-\frac{1}{2}xJ_\nu'(x)$ as an integral of (2). In each case it was noted that the function on the right hand side of the equation is a solution of the homogeneous equation obtained by replacing the right hand side by zero and, further, that the particular integral is proportional to x times the first derivative with respect to x of the function on the right hand side. These common features of the two results suggested the investigation which follows.

2. Let L and M be linear operators defined thus:

$$L(y) = y^{(n)}(x) + \sum_{r=1}^{n} p_r(x)y^{(n-r)}(x), \tag{3}$$

$$M(u) = \sum_{r=1}^{n} q_r(x)u^{(n-r)}(x). \tag{4}$$

It will be assumed that the coefficients p_r are continuous and differentiable as often as required and that the coefficients q_r are continuous.

It will be convenient to use the notation

$$(\vartheta y)(x) = xy'(x). \tag{5}$$

Let $z=u(x)$ be a solution of the homogeneous equation

$$L(z) = 0. \tag{6}$$

We consider the conditions under which the inhomogeneous equation

$$L(y) = M(u) \tag{7}$$

has a particular integral $y=\mu\vartheta u$ for some constant μ.

Using the relation $p(\vartheta u)^{(m)}=(\vartheta+m)(pu^{(m)})-xp'u^{(m)}$, $m=0, 1, 2, \cdots$, and defining the linear differential expression $K_1(u)$ by the equation

$$K_1(u) = -\sum_{r=1}^{n} x^{1-r}(x^r p_r)' u^{(n-r)}, \tag{8}$$

we can show easily that $L(\vartheta u)=(\vartheta+n)L(u)+K_1(u)$ and hence, since $L(u)=0$,

$$L(\vartheta u) = K_1(u). \tag{9}$$

It follows that the condition that $y=\mu\vartheta u$ be a particular integral of (7) is that $\mu K_1(u)=M(u)$; i.e.,

$$\sum_{r=1}^{n} \left\{q_r + \mu x^{1-r}(x^r p_r)'\right\} u^{(n-r)} = 0. \tag{10}$$

Two separate questions can now be discussed. First, we can seek the conditions under which there exists a constant μ such that $\mu\vartheta u$ is a solution of (7) for *every* solution u of (6). Alternatively, we can consider the problem of finding a value of μ such that $\mu\vartheta u$ is a solution of (7) when u is a specified solution of (6).

3. We examine first the case in which $\mu\vartheta u$ is to be a solution of (7) for all solutions u of the homogeneous equation (6) so that condition (10) must be satisfied for all such functions. Equation (10) is a linear differential equation for u of order $n-1$ which can be satisfied by all n independent solutions of the nth order equation (6) only if the coefficients of u and its derivatives in (10) all vanish. This means that there must exist a constant μ such that

$$\mu x^{1-r}(x^r p_r)' + q_r = 0, \qquad r = 1, 2, \cdots n. \tag{11}$$

Conversely, (9) may be used to show that if the coefficients p_r satisfy (11) and $L(u)=0$, then $L(\mu\vartheta u)=M(u)$. Hence we have:

THEOREM 1. *Equation* (7) *has a solution of the form* $y=\lambda xu'$ (*where* λ *is a constant*) *for every solution* $z=u$ *of* (6) *if and only if there exists a constant* μ *such that*

$$\mu x^{1-r}(x^r p_r)' + q_r = 0, \qquad r = 1, 2, \cdots, n;$$

and, moreover, $\lambda=\mu$ *so that equation* (7) *has the solution* $y=\mu xu'$.

As an example of Theorem 1 consider the case in which $q_1 = q_2 = \cdots = q_{n-1} = 0$, $q_n = \lambda$, where λ is constant. Then $L(y) = \lambda u$ has solution $y = \mu x u'$ for every u such that $L(u) = 0$ provided $p_i = a_i/x^i$, $i = 1, 2, \cdots, n-1$,

$$p_n = \frac{a_n}{x^n} - \frac{\lambda}{\mu n},$$

where a_j are arbitrary constants.

The two equations (1) and (2) which prompted this discussion are examples of the second order equation of this type and it can now be seen that in (2) the right hand side of the equation could be any solution of Bessel's equation of order ν.

A more restricted example still is obtained by taking $\lambda = 0$. We then have $y = xu'$ a solution of $L(y) = 0$ whenever $y = u$ is a solution. Clearly the condition for this is that $p_r = a_r x^{-r}$ for $r = 1, 2, \cdots, n$ so that we have an Euler equation of the form

$$x^n y^{(n)} + a_1 x^{n-1} y^{(n-1)} + \cdots + a_n y = 0. \tag{12}$$

4. We now turn to the problem of determining whether μ can be found so that $y = \mu x u'$ is a solution of (7) for a specified solution u of (6). All that is needed here is that this particular function u and the coefficients p_r and q_r should be such that condition (11) is satisfied i.e. that $\mu K_1(u) = M(u)$ for some value of μ. There are three possibilities:

(a) that condition (11) is satisfied for some μ and a particular integral $y = \mu x u'$ is found;

(b) that there is no μ for which $\mu K_1(u) = M(u)$ so that no integral of the desired form is found;

(c) that, as a special case of (b), $K_1(u) = 0$ while $M(u) \neq 0$. Then (9) gives $L(\vartheta u) = 0$ so that xu' and u are both solutions of (6) and the question now arises whether, with this additional information, a particular integral of a different form can be found.

To discuss this extended problem it is necessary to prove:

THEOREM 2. *If, for* $s \geqq 1$, $L(u) = L(xu') = \cdots = L(x^{s-1}u^{(s-1)}) = 0$, *then* $L(x^s u^{(s)}) = K_s(u)$, *where*

$$K_s(u) = (-1)^s \sum_{r=1}^{n} x^{s-r}(x^r p_r)^{(s)} u^{(n-r)}. \tag{13}$$

LEMMA 1. *If, for* $s \geqq 1$, $L(u) = L(xu') = \cdots = L(x^{s-1}u^{(s-1)}) = 0$, *then*

(a) $\qquad L(x^s u^{(s)}) = L(\vartheta^s u),$

(b) $\qquad L(u) = L(\vartheta u) = \cdots = L(\vartheta^{s-1} u) = 0.$

We can prove (a) by induction and (b) then follows by repeated application of (a).

LEMMA 2. *If, for* $s \geqq 1$, $L(u)=L(\vartheta u)= \cdots =L(\vartheta^{s-1}u)=0$, *then*

$$L(\vartheta^s u) = K_s(u).$$

This is also proved by induction, the case $s=1$ having been established earlier (equation (9)).

Theorem 2 follows at once from the two lemmas.

The theorem may be applied, in particular, when the coefficients p_r are all of the form $a_r x^{-r}$ where the a_r are constants. In this case, it is clear from (13) that $K_s(u)=0$ for $s=1, 2, 3, \cdots$ and Theorem 2 shows that if $y=u$ is a solution of $L(y)=0$ so also are $y=x^s u^{(s)}$ for $s=1, 2, 3, \cdots$. But we are now dealing with an Euler equation such as (12) which has solutions of the type

$$u = x^\alpha(A_0 + A_1 \log x + \cdots + A_p(\log x)^p).$$

It is easily verified that for any positive integer s, $x^s u^{(s)}$ is of the same form as u (with different coefficients which may be all zero) and so indeed is a solution of the equation.

To find a particular integral of (7), knowing that $L(u)=0$, we now evaluate $K_1(u)$, $K_2(u)$, $\cdots$ until the first nonzero expression is obtained, say $K_s(u) =L(x^s u^{(s)})$. If then μ can be found so that $\mu K_s(u)=M(u)$, it has been shown that $y=\mu x^s u^{(s)}$ is a particular integral of (7) and, incidentally, that $y=u$, xu', x^2u'', $\cdots$, $x^{s-1}u^{(s-1)}$ are all integrals of (6). If $\mu K_s(u) \neq M(u)$ for any μ, the method fails.

We may have started with a different member of this family of integrals of (6), say $v=x^t u^{(t)}$ where $1 \leqq t \leqq s-1$. In this case, because

$$x^{s-t}v^{(s-t)} = x^s u^{(s)} + (s - t)tx^{s-1}u^{(s-1)} + \cdots,$$

we should have obtained $K_{s-t}(v)=L(x^{s-t}v^{(s-t)})=L(x^s u^{(s)})$ as the first nonzero expression to be compared with the right hand side of the inhomogeneous equation.

As examples, we may consider the equations

$$\text{(i)} \quad y'' + 4xy' + 2(1 + 2x^2)y = e^{-x^2} \quad \text{(cf. [1]),}$$

$$\text{(ii)} \quad y'' + y'\left(4x - \frac{1}{x}\right) + 4x^2 y = x^2 e^{-x^2}.$$

For each of these $u=e^{-x^2}$ is a solution of the homogeneous equation. For (i), $K_1(u) = -4e^{-x^2}$ so that a particular integral is $-\frac{1}{4}xu'=\frac{1}{2}x^2e^{-x^2}$. For (ii), $K_1(u)=0$ so that xu' and hence $x^2e^{-x^2}$ is a solution of the homogeneous equation. $K_2(u) =32x^2e^{-x^2}$ so that a particular integral is $\frac{1}{32}x^2u''=(\frac{1}{8}x^4-\frac{1}{16}x^2)e^{-x^2}$. Clearly a simpler particular integral for (ii) is $y=\frac{1}{8}x^4e^{-x^2}$.

The method can be illustrated further in the case in which the coefficients p_r and q_r are all constants and equal respectively to P_r and Q_r. In this case we

have $L(u) = f(D)u$, where $f(D)$ is the polynomial given by

$$f(D) = D^n + \sum_{r=1}^{n} P_r D^{n-r}.$$

There are now solutions of the homogeneous equation $L(y) = 0$ of the form $y = e^{ax}$ where $f(a) = 0$.

To apply the theory to this case of constant coefficients we need:

THEOREM 3. *When* $L(y) = f(D)y$ *and* $u = e^{ax}$, *then, for* $s \geqq 1$,

$$K_s(u) = e^{ax} \sum_{t=0}^{s} (-1)^{s+r} \binom{s}{t} (n - t)_s a^t f^{(t)}(a),$$

where $(n-t)_s = (n-t)(n-t-1) \cdots (n-s+1)$, $t \leqq s-1$; $(n-s)_s = 1$; *and* $\binom{s}{t}$ *is a binomial coefficient.*

LEMMA 1. *If* $L(D)y = f(D)y$ *and* $u = e^{ax}$, *then, for* $s \geqq 1$,

$$K_s(u) = (-1)^s e^{ax} A_s(a),$$

where $A_s(a) = \sum_{r=1}^{n} r(r-1) \cdots (r-s+1) P_r a^{n-r}$.

This follows from substituting $u = e^{ax}$, $p_r = P_r$ in the expression (13) for $K_s(u)$.

LEMMA 2. *For* $s \geqq 1$, $(n-s)A_s(a) - aA_s'(a) = A_{s+1}(a)$.

LEMMA 3. *For* $s \geqq 1$,

$$A_s(a) = \sum_{t=0}^{s} (-1)^t \binom{s}{t} (n - t)_s a^t f^{(t)}(a).$$

This is proved by induction, using Lemma 2. Theorem 3 follows from Lemmas 1 and 3.

Now suppose that $f(a) = f'(a) = \cdots f^{(s-1)}(a) = 0$, $f^{(s)}(a) \neq 0$, and $u = e^{ax}$. Then $L(u) = 0$ and from Theorem 3, $K_1(u) = K_2(u) = \cdots K_{s-1}(u) = 0$, and $K_s(u) = e^{ax} a^s f^{(s)}(a)$. When $u = e^{ax}$ and $q_r = Q_r$, $M(u) = k(a)e^{ax}$ where $k(a) = \sum_{r=1}^{n} Q_r a^{n-r}$.

In this case then, we do have $\mu K_s(u) = M(u)$ provided

$$\mu a^s f^{(s)}(a) = k(a). \tag{14}$$

A particular integral of the equation

$$f(D)y = ke^{ax} \tag{15}$$

when $f(a) = f'(a) = \cdots = f^{(s-1)}(a) = 0$, $f^{(s)}(a) \neq 0$, is therefore $y = \mu x^s u^{(s)}$ where $u = e^{ax}$ and μ is given by (14), i.e., we have the familiar result that $y = kx^s e^{ax}/f^{(s)}(a)$ is a particular integral of (15).

5. A similar investigation can be conducted into the conditions under which equation (7) has a particular integral of the form $\lambda u'$ where λ is a constant and $z=u$ is a solution of equation (6). Theorems similar to Theorems 1 and 2 are in this case rather more easily derived:

THEOREM 1 (a). *Equation* (7) *has a solution of the form* $y=\lambda u'$ (*where* λ *is a constant*) *for every solution* $z=u$ *of equation* (6) *if and only if there exists a constant* μ *such that*

$$\mu p_r' + q_r = 0, \qquad r = 1, 2, \cdots n;$$

and, moreover, $\lambda=\mu$, *so that equation* (7) *has the solution* $y=\mu u'$.

THEOREM 2 (a). *If, for* $s\geqq 1$, $L(u)=L(u')= \cdots L(u^{(s-1)})=0$, *then* $L(u^{(s)}) = \overline{K}_s(u)$, *where* $\overline{K}_s(u)=(-1)^s\sum_{r=1}^{n} p_r^{(s)}u^{(n-r)}$.

Either Theorems 1 and 2 or Theorems 1(a) and 2(a) can be used to find the particular integrals in the cases considered so far in this paper, for the change of variable $x=e^t$ changes xdu/dx into du/dt and changes equation (7) into an equation which is of the same form as (7) but with t as independent variable. The discussion in terms of the variable x has been preferred because in equations (1) and (2) and in the examples discussed in section 4, the variable x is obviously a more natural one to use than t in the right hand sides of the equations.

On the other hand, the Euler equation (12) is more easily discussed after the variable has been changed to t and Theorems 1(a) and 2(a) open up a new range of examples. For instance, if all the coefficients $p_r(x)$ are linear functions a_rx+b_r and the coefficients $q_r(x)$ are the constants $-\mu a_r$, Theorem 1(a) shows that the equation

$$y^{(n)} + \sum_{r=1}^{n} (a_rx + b_r)y^{(n-r)} = -\mu \sum_{r=1}^{n} a_r u^{(n-r)}$$

has solution $y=\mu u'$ where $y=u$ is any solution of the homogeneous equation.

I have had the benefit of useful discussions of this problem with several of my colleagues and I wish to express my thanks particularly to Mr. W. A. Coppel, Dr. J. B. Miller and Dr. M. F. Newman. I am indebted also to the referee for a suggestion which led to the inclusion of section 5.

Reference

1. E. Kamke, Differentialgleichungen, Vol. 1, Akademische Verlagsgesellschaft, Leipzig, 1943. (2.47, p. 415).

ORDINARY DIFFERENTIAL EQUATIONS OF THE FOURTH ORDER WITH ORTHOGONAL POLYNOMIAL SOLUTIONS

JOHN W. MEUX, Midwestern University

Introduction. The late W. C. Brenke [1] developed a pair of conditions for orthogonality of polynomial solutions of equations of the form

$$P(x)y_n'' + Q(x)y_n' + \lambda_n y_n = 0, \tag{1}$$

where $P(x)$ and $Q(x)$ are polynomials in x and λ_n is a polynomial in n, $\lambda_0 = 0$.

In order that equation (1) have polynomial solutions of the form

$$y_n = \sum_{j=0}^{n} a_{jn}x^{n-j}, \qquad a_{0n} \neq 0. \tag{2}$$

Brenke showed the necessity of the conditions

$$\begin{aligned} P(x) &= \alpha x^2 + \beta x + \gamma, \\ Q(x) &= -\left(x + \frac{a_{11}}{a_{01}}\right), \\ \lambda_n &= n - n(n-1)\alpha. \end{aligned} \tag{3}$$

Furthermore, he showed that the set $\{y_n(x)\}$, $n = 0, 1, 2, \cdots$ of solutions of equation (1) is orthogonal over (a, b) with respect to a weight function $W(x)$ if

$$\begin{aligned} W(x)P(x) &= 0 \text{ at } x = a \text{ and at } x = b, \\ [W(x)P(x)]' &\equiv W(x)Q(x). \end{aligned} \tag{4}$$

If equation (1) is multiplied by a weight function $W(x)$ and if the identity in equation (4) is applied then equation (1) can be written in the self-adjoint form

$$(WPy_n')' + \lambda_n W y_n = 0. \tag{5}$$

Brenke derives weight functions for each of the three cases of the finite interval, the semi-infinite interval and the infinite interval. These weight functions are the same as those for the classical orthogonal polynomials, which are classified (in [2]) in the following manner:

Jacobi polynomials, $W(x) = (1-x)^\alpha(x+1)^\beta$, $\alpha > -1$, $\beta > -1$ on $(-1, 1)$,
Laguerre polynomials, $W(x) = x^\alpha e^{-x}$, $\alpha > -1$ on $(0, \infty)$,
Hermite polynomials, $W(x) = e^{-x^2/2}$ on $(-\infty, \infty)$.
There are two well-known subcases of Jacobi polynomials, namely:
Legendre polynomials, $W(x) = 1$ on $(-1, 1)$,
Tchebichef polynomials, $W(x) = (1-x^2)^{-1/2}$ on $(-1, 1)$.

An interval (a, b) can be transformed into $(-1, 1)$ and half-lines (a, ∞) or $(-\infty, b)$ into $(0, \infty)$ by simple substitutions (see [3]) so that the basic intervals under consideration are those of the classical polynomials given above.

The problem of this paper is to find a class of fourth order differential equations having these orthogonal polynomials as solutions.

Polynomial solutions. Consider the equation

$$P(x)y_n^{\mathrm{iv}} + Q(x)y_n''' + R(x)y_n'' + S(x)y_n' + \lambda_n y_n = 0. \tag{6}$$

It can be shown that for this equation to have polynomial solutions of the form

$$y_n = \sum_{j=0}^{n} a_{jn}x^{n-j}, \qquad a_{0n} \neq 0, \tag{7}$$

it is necessary that $P(x)$ be a polynomial of at most degree four, $Q(x)$ be a polynomial of at most degree three, $R(x)$ be a polynomial of at most degree two, $S(x)$ be a polynomial of degree one, and λ_n be a polynomial in n. No loss of generality occurs if the leading coefficient of $P(x)$ is chosen as unity. Specifically,

$$\begin{aligned} P(x) &= A_1x^4 + A_2x^3 + A_3x^2 + A_4x + A_5, \\ Q(x) &= B_1x^3 + B_2x^2 + B_3x + B_4, \\ R(x) &= C_1x^2 + C_2x + C_3, \\ S(x) &= D_1x + D_2,\ D_1 \neq 0 \\ \lambda_n &= -n\{D_1 + (n-1)\{C_1 + (n-2)[B_1 + A_1(n-3)]\}\}. \end{aligned} \tag{8}$$

Orthogonality. The following theorems will establish sufficient conditions for the orthogonality of the solution set (7) of equation (6), over (a, b) with respect to a suitably chosen weight function $W(x)$.

THEOREM I. *If $W(x)$ and its first three derivatives are continuous and differentiable in (a, b) and if*

$$\text{(i) } 2(WP)' \equiv WQ, \qquad \text{(ii) } WS \equiv (WR)' - (WP)''', \tag{9}$$

then each equation of the family (6), after multiplication by $W(x)$, is self-adjoint.

Proof. Multiplication of equation (6) by $W(x)$ and the application of conditions (i) and (ii) yields

$$WPy_n^{\mathrm{iv}} + 2(WP)'y_n''' + WRy_n'' + [(WR)' - (WP)''']y_n' + \lambda_n Wy_n = 0. \tag{10}$$

The addition, and subsequent subtraction, of $(WP)''y_n''$ and grouping gives

$$\begin{aligned} &[WPy_n^{\mathrm{iv}} + 2(WP)'y_n''' + (WP)''y_n''] \\ &\qquad + \{[WR - (WP)'']y_n'' + [(WR)' - (WP)''']y_n'\} + \lambda_n Wy_n = 0. \end{aligned} \tag{11}$$

Application of Leibniz's rule for derivatives of higher order gives

$$(WPy_n'')'' + \{[WR - (WP)'']y_n'\}' + \lambda_n W y_n = 0, \tag{12}$$

thus rendering equation (6) self-adjoint.

THEOREM II. *If the conditions of Theorem* I *hold and if*

$$\begin{aligned} &\text{(iii)} && WP = 0 && \text{at} \quad x = a \quad \text{and at} \quad x = b, \\ &\text{(iv)} && (WP)' = 0 && \text{at} \quad x = a \quad \text{and at} \quad x = b, \\ &\text{(v)} && WR - (WP)'' = 0 && \text{at} \quad x = a \quad \text{and at} \quad x = b, \end{aligned} \tag{13}$$

hold, and $\lambda_n \neq \lambda_m$, *then respective solutions of* (6) *are orthogonal.*

Proof. Consider

$$(WPy_n'')'' + \{[(WR) - (WP)'']y_n'\}' + \lambda_n W y_n = 0, \tag{14}$$

$$(WPy_m'')'' + \{[(WR) - (WP)'']y_m'\}' + \lambda_m W y_m = 0. \tag{15}$$

Multiplication of equations (14) and (15) by y_m and $-y_n$ respectively, addition of the resulting equations and subsequent integration over (a, b) with respect to x yields

$$\begin{aligned} &\int_a^b y_m\{(WPy_n'')'' + \{[(WR) - (WP)'']y_n'\}'\}\, dx - \int_a^b y_n\{(WPy_m'')'' \\ &\qquad + \{[(WR) - (WP)'']y_m'\}'\}\, dx + (\lambda_n - \lambda_m)\int_a^b W y_n y_m\, dx = 0. \end{aligned} \tag{16}$$

Let $I_m = \int_a^b y_m\{(WPy_n'')'' + \{[(WR) - (WP)'']y_n'\}'\}dx$. Integration by parts and application of boundary conditions (iii), (iv) and (v) yields

$$I_m = -\int_a^b y_m'(WPy_n'')'\, dx - \int_a^b [(WR) - (WP)'']y_n' y_m'\, dx. \tag{17}$$

Similarly, if

$$I_n = \int_a^b y_n\{(WPy_m'')'' + \{[(WR) - (WP)'']y_m'\}'\}\, dx,$$

then

$$I_n = -\int_a^b y_n'(WPy_m'')'dx - \int_a^b [(WR) - (WP)'']y_m' y_n' dx. \tag{18}$$

Thus,

$$I_m - I_n = \int_a^b y_n'(WPy_m'')'\, dx - \int_a^b y_m'(WPy_n'')'\, dx. \tag{19}$$

Let $I_{\bar{n}}=\int_a^b y_n'(WPy_m'')'dx$. Again, integration by parts and the application of boundary condition (iii) gives

$$I_{\bar{n}} = -\int_a^b WPy_m''y_n''\,dx. \tag{20}$$

Similarly, if

$$I_m = \int_a^b y_m'(WPy_n'')'\,dx,$$

then

$$I_{\bar{m}} = -\int_a^b WPy_n''y_m''\,dx. \tag{21}$$

Thus,

$$I_{\bar{n}} - I_{\bar{m}} = 0 = I_m - I_n, \tag{22}$$

which implies that

$$(\lambda_n - \lambda_m)\int_a^b Wy_ny_m\,dx = 0. \tag{23}$$

If the eigenvalues λ_n and λ_m are not equal, then the integral part of equation (23) is zero and, therefore, the respective solutions of equation (6) are orthogonal in (a, b) with respect to $W(x)$.

The finite interval. Consider the Jacobi polynomials obtained by choosing

$$W(x) = (1-x)^\alpha(x+1)^\beta,\ \alpha > -1,\ \beta > -1 \text{ on } (-1, 1).$$

THEOREM III. *If the conditions of Theorems* I *and* II *hold, if* $a=-1$, $b=1$, *and if* $W(x)=(1-x)^\alpha(x+1)^\beta$, $\alpha>-1$, $\beta>-1$, *then the coefficient* $P(x)$ *in equation* (12) *is*

$$P(x) = (1-x)^2(x+1)^2. \tag{24}$$

Proof. Assume that $P(-1)\neq 0$. Application of the self-adjoint condition (i) yields

$$\begin{aligned}(1-x)^\alpha(x+1)^\beta Q \equiv\ & 2(1-x)^\alpha(x+1)^\beta P' \\ & + 2P[\beta(1-x)^\alpha(x+1)^{\beta-1} - \alpha(1-x)^{\alpha-1}(x+1)^\beta],\end{aligned} \tag{25}$$

or

$$Q \equiv 2P' + 2P[\beta(x+1)^{-1} - \alpha(1-x)^{-1}]. \tag{26}$$

The polynomial nature of Q requires that $1-x$ be a factor of $P(x)$. This contradicts the assumption that $P(-1)\neq 0$; hence $P(-1)=0$. Similarly, it can be shown that $P(1)=0$.

Assume that $P(x)$ has simple zeros of -1 and 1. Then

$$(27) \qquad P(x) = (1-x)(x+1)U(x),$$

where $U(-1) \neq 0$ and $U(1) \neq 0$. Application of the self-adjoint condition (ii) gives

$$(28) \quad (1-x)^{\alpha}(x+1)^{\beta}S \equiv [(1-x)^{\alpha}(x+1)^{\beta}R]' - [(1-x)^{\alpha+1}(x+1)^{\beta+1}U]'''.$$

Performance of the indicated differentiation and solution of equation (28) for S will show that the polynomial nature of $S(x)$ requires that both $1-x$ and $x+1$ be factors of $U(x)$. This contradicts the assumption that $U(-1) \neq 0$ and $U(1) \neq 0$, thus implying that $P(x)$ has both -1 and 1 as double zeros.

Since $P(x)$ is a polynomial of degree at most four,

$$P(x) = (1-x)^2(x+1)^2.$$

THEOREM IV. *If the conditions of Theorems* I, II *and* III *hold, then the coefficient* $R(x)$ *in equation* (12) *is a polynomial of degree two, such that* $R(-1) = 4(\beta+1)(\beta+2)$ *and* $R(1) = 4(\alpha+1)(\alpha+2)$.

Proof. Boundary condition (v) states that

$$(29) \quad (1-x)^{\alpha}(x+1)^{\beta}R - [(1-x)^{\alpha+2}(x+1)^{\beta+2}]'' = 0 \text{ at } x = -1 \text{ and at } x = 1;$$

hence

$$(30) \quad \begin{aligned} R(x) = {} & (\alpha+2)(\alpha+1)(x+1)^2 - 2(\alpha+2)(\beta+2)(1-x)(x+1) \\ & + (\beta+2)(\beta+1)(1-x)^2 \quad \text{at} \quad x = -1 \quad \text{and} \quad x = 1, \end{aligned}$$

implying that $R(-1) = 4(\beta+2)(\beta+1)$ and $R(1) = 4(\alpha+2)(\alpha+1)$.

Example of the finite interval. As in the case of the Legendre polynomials choose $\alpha = \beta = 0$, so that $W(x) = 1$. By Theorems III and IV, respectively,

$$(31) \qquad P(x) = (1-x)^2(x+1)^2$$

and

$$(32) \qquad R(x) = C_1x^2 + (8 - C_1).$$

From the self-adjoint condition (ii)

$$(33) \qquad S \equiv R' - P''' \equiv 2C_1x - 24x.$$

The choice of C_1 is arbitrary with the exceptions that C_1 must be chosen so that $\lambda_n \neq \lambda_m$, $n \neq m$, and $C_1 \neq 12$, as $S(x)$ must be a first degree polynomial. Hence, choose $C_1 = 14$, from which it is found that

$$(34) \qquad \lambda_n = -n^2(n+1)^2.$$

The differential equation (12) is thus

$$(35) \qquad [(1-x)^2(x+1)^2 y_n'']'' + [2(x^2-1)y_n']' - n^2(n+1)^2 y_n = 0.$$

The application of series techniques to equation (35) will yield precisely the Legendre polynomials, provided that the arbitrary constant is chosen so that $y_n(1)=1$.

The semi-infinite interval. Consider the Laguerre polynomials obtained by choosing $W(x)=x^\alpha e^{-x}$, $\alpha>-1$ on $(0, \infty)$. The following two theorems are stated without proof as their proofs are quite similar to those of Theorems III and IV, respectively.

THEOREM V. *If the conditions of Theorems I and II hold, if the interval is* $(0, \infty)$ *and if* $W(x)=x^\alpha e^{-x}$, $\alpha>-1$, *then the coefficient* $P(x)$ *in equation* (12) *is*

$$P(x) = x^2. \tag{36}$$

THEOREM VI. *If the conditions of Theorems I, II and V hold, then the coefficient* $R(x)$ *in equation* (12) *is*

$$R(x) = x^2 + C_2x + C_3, \tag{37}$$

where $R(0)=(\alpha+1)(\alpha+2)$.

Example of the semi-infinite interval. Let $\alpha=0$. Then, by Theorems V and VI,

$$P(x) = x^2, \tag{38}$$

and

$$R(x) = x^2 + C_2x + 2. \tag{39}$$

From the self-adjoint condition (ii) it is seen that

$$S(x) = (4 + C_2)(1 - x). \tag{40}$$

The choice of C_2 is subject to the same restrictions as in the preceding example. The choice $C_2=-5$ is suitable in this case and λ_n is found to be

$$\lambda_n = -n^2. \tag{41}$$

The differential equation (12) is

$$(x^2e^{-x}y_n'')'' - (xe^{-x}y_n')' - n^2e^{-x}y_n = 0. \tag{42}$$

Solutions of this equation are precisely the Laguerre polynomials, provided the arbitrary constant is chosen so that the leading coefficient of each polynomial is unity.

The infinite interval. By choosing $W(x)=e^{-x^2/2}$ on $(-\infty, \infty)$ we obtain the Hermite polynomials. The theorems corresponding to Theorems III and IV, respectively, are stated below.

THEOREM VII. *If the conditions of Theorems* I *and* II *hold, if the interval is* $(-\infty, \infty)$ *and if* $W(x)=e^{-x^2/2}$, *then* $P(x)$ *in equation* (12) *is*

$$P(x) = 1. \tag{43}$$

THEOREM VIII. *If the conditions of Theorems* I, II *and* VII *hold, then* $R(x)$ *in equation* (12) *is*

$$R(x) = x^2 + C_3. \tag{44}$$

Example of the infinite interval. By Theorems VII and VIII,

$$P(x) = 1, \tag{45}$$

and

$$R(x) = x^2 + C_3. \tag{46}$$

The self-adjoint condition (ii) requires

$$S(x) = -(C_3 + 1)x. \tag{47}$$

The choice of C_3 is subject to the same restrictions as in the preceding examples. A suitable choice is $C_3 = -2$, from which it is found that

$$\lambda_n = -n^2. \tag{48}$$

The differential equation (12) is

$$(e^{-x^2/2}y_n'')'' - (e^{-x^2/2}y_n')' - n^2e^{-x^2/2}y_n = 0. \tag{49}$$

Solutions of this equation are precisely the Hermite polynomials if the arbitrary constant is chosen so that the leading coefficient of each polynomial is unity.

References

1. W. C. Brenke, On polynomial solutions of a class of linear differential equations of the second order, Bull. Amer. Math. Soc., 36 (1930) 77–84.

2. Dunham Jackson, Fourier series and orthogonal polynomials, MAA Carus Monograph Number Six, Buffalo, 1941; Chapters VII, VIII, IX, X.

3. F. G. Tricomi, Vorlesungen über Orthogonalreihen, Springer, 1955; Chapter IV.

4. G. Szegö, Orthogonal Polynomials, Amer. Math. Soc. Colloq. Publ., New York 23, (1939).

ON DIFFERENTIABLE TRANSFORMATIONS IN R^n

D. E. VARBERG, Hamline University

1. Introduction. Our purpose is to give a proof and note some consequences of the following

THEOREM 1. *Let f be a mapping of an open subset D of R^n into R^n. Let E be any measurable subset of D where f is differentiable and let $J(x)$ denote the Jacobian determinant of f at x. Then the set $f(E)$ is measurable and*

$$mf(E) \leqq \int_E |J(x)| \, dx,$$

where m denotes n-dimensional Lebesgue measure.

In a recent issue of this MONTHLY [**4**], T. M. Flett has called attention to this result, but his statement and proof require that f be continuously differentiable and this severely limits its applications. The theorem as stated follows from very deep results of Federer [**2**, pp. 449–452], [**3**, pp. 426–432] and is closely related to others of Rado and Reichelderfer [**7**, pp. 363–365]. Our proof (in the spirit of Flett's paper) will avoid difficult topological notions but will require the generalization to R^n of two lemmas which appear in standard textbooks for R^1. Some consequences of Theorem 1 will be an important theorem of Sard on the image of critical points and some results on absolutely continuous mappings on R^n.

2. Proof of Theorem 1. We begin with two lemmas. For the corresponding results for R^1, see Saks [**8**, pp. 226, 271] and for a generalization of the first lemma, see [**10**, p. 248]. Sets of Lebesgue measure zero will be called null sets

LEMMA 1. *Let N be a null subset of D and suppose that f is differentiable on N. Then $mf(N)=0$.*

Proof. For each pair of positive integers j and k, let N_{jk} be the set of points x of N for which

$$(2.1) \qquad \|f(x+t) - f(x)\| \leqq j\|t\|$$

for all t such that $\|t\| \leqq 1/k$. Since f is differentiable on N, $N = \bigcup N_{jk}$ and so it will suffice to show that $mf(N_{jk})=0$. Suppose then that j and k are fixed and let $\epsilon>0$ be given. Noting that $m(N_{jk})=0$, we may choose a sequence of n-cubes $\{V_i\}$ with centers $\{v_i\}$ and sides of length $\{2b_i\}$ so that $b_i \leqq 1/nk$, $N_{jk} \subset \bigcup V_i$ and $\sum m V_i < \epsilon$. On $V_i \cap N_{jk}$, $\|v - v_i\| < nb_i \leqq 1/k$ and so by (2.1)

$$\|f(v) - f(v_i)\| \leqq j\|v - v_i\| < jnb_i.$$

Thus $f(V_i \cap N_{jk})$ is contained in an n-cube with center $f(v_i)$ and side length $2jnb_i$. Letting m^* denote exterior Lebesgue measure, we see that

$$m^*f(V_i \cap N_{jk}) \leqq (2jnb_i)^n = (jn)^n mV_i$$

and therefore

$$m^*f(N_{jk}) \leqq \sum_i m^*f(V_i \cap N_{jk}) \leqq (jn)^n \sum_i mV_i \leqq (jn)^n \epsilon.$$

But ϵ was an arbitrary positive number. Thus $mf(N_{jk}) = 0$.

LEMMA 2. *Let E be any subset of D where f is differentiable and $|J(x)| \leqq K$. Then*

$$m^*f(E) \leqq Km^*E, \tag{2.2}$$

where m^* denotes exterior Lebesgue measure.

Proof. We may suppose that $m^*E < \infty$. Let $\epsilon > 0$ be given and choose an open subset A of D such that $E \subset A$ and $mA \leqq m^*E + \epsilon$. For each x in E, there exists a closed oriented (i.e., sides parallel to axes) n-cube E_x with center x such that $E_x \subset A$ and

$$m^*f(E_x) \leqq (K + \epsilon)mE_x. \tag{2.3}$$

To see this, we note that since f is differentiable at x

$$\|f(y) - f(x) - f'(x)(y - x)\| = \eta(y)\|y - x\|,$$

where $\eta(y) \to 0$ as $y \to x$ and $f'(x)$ denotes the derivative of f at x in the sense of Dieudonné [**1**, p. 143]. Hence, just as in the proof of Lemma 4 of [**4**], for any oriented n-cube E_x centered at x

$$m^*f(E_x) \leqq mE_x[\,|J(x)| + A(n)(\|f'(x)\| + \eta)^{n-1}\eta],$$

where $A(n)$ is constant except for its dependence on n and

$$\eta = \sup_{y \in Ex} \eta(y).$$

Using now the fact that $\eta(y) \to 0$ as $y \to x$ and that $|J(x)| \leqq K$, we see that (2.3) holds for all sufficiently small oriented n-cubes about x.

Next consider the collection $\mathfrak{F}_x$ of all oriented n-cubes E_x, contained in A, centered at x and satisfying (2.3). Finally let $\mathfrak{F}$ be the family of all cubes so obtained as x ranges over E. This family forms a Vitali covering of E and hence [**8**, p. 109] $\mathfrak{F}$ contains a denumerable or finite collection of pairwise disjoint cubes $E_1, E_2, \cdots$ such that $m(E - \cup E_i) = 0$. Letting $N = E - \cup E_i$, we observe that

$$m^*f(E) \leqq m^*f(\cup E_i) + m^*f(N).$$

By Lemma 1, $m^*f(N)=0$. On the other hand,

$$\begin{aligned} m^*f(\cup E_i) &\leqq \sum m^*f(E_i) \leqq \sum (K+\epsilon)mE_i \\ &= (K+\epsilon)m(\cup E_i) \leqq (K+\epsilon)mA \\ &\leqq (K+\epsilon)(m^*E+\epsilon). \end{aligned}$$

Since $\epsilon>0$ was arbitrary, the result follows.

We proceed to the proof of Theorem 1. We may assume that $mE<\infty$ for if not it is only necessary to write E as the disjoint union of such sets and work with each of these. Our first project is to show that $f(E)$ is measurable. But this follows in a standard way from the continuity of f on E and the fact (Lemma 1) that f maps null subsets of E into null sets (see [**6**], p. 248).

Next for any $\epsilon>0$ and positive integer k, let

$$E_k = \{x \in E \mid (k-1)\epsilon \leqq |J(x)| < k\epsilon\}.$$

Then E_k is measurable and hence so is $f(E_k)$ and thus

$$\begin{aligned} mf(E) &\leqq \sum mf(E_k) \leqq \sum k\epsilon mE_k && \text{(by Lemma 2)} \\ &= \sum (k-1)\epsilon mE_k + \epsilon \sum mE_k \leqq \int_E |J(x)|\, dx + \epsilon mE. \end{aligned}$$

The result now follows from the arbitrariness of ϵ.

3. Some consequences. We obtain easily the promised Theorem of Sard [**9**], [**10**, p. 254].

THEOREM 2. *Let E be any subset of D on which f is differentiable and $J(x)=0$. Then $mf(E)=0$.*

Proof. If E is measurable, we may apply Theorem 1. In any case, the result follows immediately from Lemma 2 with $K=0$.

Several theorems about absolutely continuous mappings are trivial consequences of Theorem 1. For completeness we give the definition of absolute continuity (Banach sense). First for any open set D in R^n, a system of intervals $I_1, \cdots, I_m$ is termed *admissible* for D if $\operatorname{int}(I_k)\subset D$, $k=1,\cdots,m$, and $\operatorname{int}(I_j)\cap\operatorname{int}(I_k)$ is empty for $j\neq k$. A bounded continuous mapping f of an open subset D of R^n into R^n is said to be *absolutely continuous on D* if given any $\epsilon>0$ there exists a corresponding $\delta>0$ such that $\sum m^*f(I_k)<\epsilon$ for every admissible system of intervals for D with $\sum mI_k<\delta$ (cf. [**7**] p. 283).

THEOREM 3. *If on D, f is bounded and continuous, almost everywhere differentiable, and maps null sets into null sets, and if $J(x)$ is integrable, then f is absolutely continuous on D.*

Proof. Let $I_1, \cdots, I_m$ be an admissible system of intervals for D and let E_k be the subset of I_k where f is differentiable. Then $f(E_k)$ is measurable (Theorem 1) and hence so is $f(I_k)$. Moreover,

$$\sum mf(I_k) = \sum mf(E_k) \leqq \sum \int_{E_k} | J(x) | \, dx = \sum \int_{I_k} | J(x) | \, dx,$$

where the inequality follows from Theorem 1. Since $J(x)$ is integrable on D, the last expression approaches zero as $\sum mI_k$ tends to zero by a well-known property of the Lebesgue integral.

THEOREM 4 (cf. [5] p. 183). *If on D, f is bounded and continuous, differentiable for all but a finite or denumerable subset and if $J(x)$ is integrable, then f is absolutely continuous on D.*

For the proof we need only note that f maps null sets into null sets and then apply Theorem 3. A corollary is

THEOREM 5 (cf. [6] p. 266). *If on D, f is bounded and differentiable and $J(x)$ is integrable, then f is absolutely continuous on D.*

Acknowledgement. The author would like to thank the referee for helpful comments and for calling his attention to several references.

This research was supported by the Air Force Office of Scientific Research.

References

1. J. Dieudonné, Foundations of modern analysis, Academic Press, New York, 1960.
2. H. Federer, Surface area II, Trans. Amer. Math. Soc., 55 (1944), 438–456.
3. ———, Curvature measures, *ibid.*, 93 (1959) 418–491.
4. T. M. Flett, On transformations in R^n and a theorem of Sard, this MONTHLY, 71 (1964) 623–629.
5. H. Kestelman, Modern theories of integration, Oxford, 1937.
6. I. P. Natanson, Theory of functions of a real variable (translated from Russian), Rev. ed., Ungar, New York, 1961.
7. T. Rado and P. V. Reichelderfer, Continuous transformations in analysis, Springer-Verlag, Berlin, 1955.
8. S. Saks, Theory of the integral, Monografie Matematyczne, Warszawa, 1937.
9. A. Sard, The measure of the critical values of differentiable maps, Bull. Amer. Math. Soc., 48 (1942) 883–890.
10. ———, Images of critical sets, Ann. Math., 68 (1958) 247–259.

THEOREM 5. *Let f be integrable on* S^{n-1}, *let* $\{S_m^k\}$ $(m=0, 1, \cdots; k=0, 1, \cdots, h_m)$ *be an orthonormal set of spherical harmonics,* S_m^k *of degree k, and* $h_m=\dim(SH^m)$. *Let* $a_{mk}=\int_{S^{n-1}} fS_m^k$. *Then there is a* C^∞ *function g,* $g=f$ *a.e., if and only if* $a_{mk}=0(m^{-r})$ *for every r.*

Proof. If g exists, then

$$|a_{mk}| = m^{-r}(m+n-2)^{-r}\left|\int g(\Delta_s)^r S_m^k\right| = m^{-r}(m+n-2)^{-r}\left|\int (\Delta_s^r g)S_m^k\right|$$

$$\leqq m^{-r}(m+n-2)^{-r}\left(\int |\Delta_s^r g|^2\right)^{1/2}.$$

Conversely, if $a_{mk}=0(m^{-r})$ for every r, then by Theorems 2 and 4, $\sum a_{mk}S_m^k$ converges to a C^∞ function g. Since $\int_{S^{n-1}}(g-f)S_m^k=0$ for all m and k, we have from Theorem 1(c) that $\int_{S^{n-1}}(g-f)p=0$ for every polynomial p. Standard density results yield $f-g=0$ a.e.

The preparation of this paper was supported in part by an NSF grant at Brandeis University.

References

1. A. Erdelyi, ed., Higher Transcendental Functions, vol. 2, New York, 1953.

2. A. P. Calderón and A. Zygmund, Singular integral operators and differential equations, Amer. J. Math., 79 (1957) 901–921.

A CHARACTERIZATION OF THE EXPONENTIAL SERIES

J. D. BUCKHOLTZ, University of Kentucky

For all sufficiently large positive integers n, K. S. K. Iyengar [4] has shown that the disc $|z| \leqq ne^{-2}$ contains no zero of $\sum_{p=0}^{n} z^p/p!$, the nth partial sum of the power series for e^z. The main purpose of this note is to show that this fact characterizes the exponential series.

THEOREM 1. *Suppose that* $\sum_{p=0}^{\infty} a_p z^p$ *is a power series. The following two statements are equivalent:*

(i) *There is a positive number c such that for every positive integer n,* $\sum_{p=0}^{n} a_p z^p$ *has no zero in the disc* $|z| \leqq nc$,

(ii) $a_0 \neq 0$, *and* $\sum_{p=0}^{\infty} a_p z^p$ *is the power series for* $a_0 \exp(a_1 z/a_0)$.

Proof. It follows from [4] that (ii) implies (i). Suppose now that (i) is true. Then

$$nc < |a_0/a_n|^{1/n}, \tag{1}$$

since the quantity on the right is the geometric mean of the moduli of zeros of $\sum_{p=0}^{n} a_p z^p$. Using (1), we have

$$\limsup_{n\to\infty} \frac{\log n}{n \log |1/a_n|} \leqq 1.$$

Therefore $\sum_{p=0}^{\infty} a_p z^p$ is an entire function of order 1 or less. That this function has no zeros follows from (i) and Hurwitz's theorem [1, p. 171]. Making use of the Hadamard factorization theorem, one obtains (ii).

Interpreted freely, Theorem 1 asserts that the zeros of partial sums of the exponential series have larger moduli than those of the partial sums of any other power series.

In studying the distribution of zeros of partial sums of the exponential series, it is convenient to "scale down" by a factor of n and consider instead the zeros of the polynomials

$$P_n(z) = \sum_{p=0}^{n} \frac{(nz)^p}{p!}, \qquad n = 1, 2, 3, \cdots.$$

G. Szegö [6] has shown that as $n \to \infty$ the zeros of $P_n(z)$ cluster along the simple closed curve Γ given by

$$\Gamma: |ze^{1-z}| \leqq 1, \qquad |z| \leqq 1,$$

and that the proportion of zeros clustering along a given arc of Γ is asymptotically equal to the change in $(2\pi)^{-1} \arg (ze^{1-z})$ along the arc.

We now prove that all the zeros of $P_n(z)$ lie on the same side of Γ, and obtain an upper bound for their distances from Γ. The following theorem improves on the result of Iyengar [4], and, in view of [6], is clearly "best possible."

THEOREM 2. *For every positive integer n neither Γ nor the bounded region it encloses contains a zero of $P_n(z)$.*

Proof. Suppose that $|ze^{1-z}| \leqq 1$ and $|z| \leqq 1$. We shall prove that $P_n(z) \neq 0$ by showing that $|1 - e^{-nz}P_n(z)| < 1$. For this purpose, we observe that

$$\begin{aligned} |1 - e^{-nz}P_n(z)| &= \left| (ze^{1-z})^n \sum_{p=n+1}^{\infty} \frac{n^p e^{-n} z^{p-n}}{p!} \right| \\ &\leqq \sum_{p=n+1}^{\infty} \frac{n^p e^{-n}}{p!} \\ &= 1 - e^{-n}P_n(1) < 1. \end{aligned}$$

This completes the proof. A similar argument can be used to show that all the zeros of the functions $R_n(z) = e^{nz} - P_n(z)$ lie on the left side of the open curve Γ_1 given by $|ze^{1-z}| = 1$, $|z| \geqq 1$.

THEOREM 3. *For every positive integer n all the zeros of $P_n(z)$ lie within a distance $2e/n^{\frac{1}{2}}$ of Γ.*

Proof. The zeros of $P_n(z)$ are the same as those of the function $T_n(z)$ defined by

$$P_n(z) = \frac{(nz)^n}{n!} T_n(z).$$

From [2, Lemma 1] we have

$$T_n(z) = \frac{z}{z-1}\left[1 + \frac{T_n'(z)}{n}\right], \qquad z \neq 0, 1, \tag{2}$$

and also, that the inequality

$$\left| T_n(z) \right| < 2en^{\frac{1}{2}} \tag{3}$$

holds for all z not in the bounded region enclosed by Γ. For points z whose distance from this region is δ or more, we have $|T_n'(z)| < 2en^{\frac{1}{2}}/\delta$ from the Cauchy inequality for derivatives. For $\delta = 2e/n^{\frac{1}{2}}$, (2) implies that $T_n(z) \neq 0$, and therefore that $P_n(z) \neq 0$.

An identical result holds for the zeros (other than $z=0$) of $R_n(z)$ and the curve Γ_1. In this case one makes use of the function $S_n(z)$ defined in [2], and replaces (2) and (3) by corresponding properties [2, Lemma 1] of $S_n(z)$.

References

1. L. V. Ahlfors, Complex analysis, McGraw-Hill, New York, 1953.
2. J. D. Buckholtz, Concerning an approximation of Copson, Proc. Amer. Math. Soc., 14 (1963) 564–568.
3. ———, Power series whose sections have zeros of large modulus, Trans. Amer. Math. Soc., 117 (1965) 157–166.
4. K. S. K. Iyengar, A note on the zeros of $\sum_0^n z^r/r! = 0$, Math. Student, 6 (1938) 77–78.
5. P. Rosenbloom, Distribution of zeros of polynomials, Lectures on functions of a complex variable, Univ. of Michigan Press, Ann Arbor, 1955, 265–285.
6. G. Szegö, Über eine Eigenschaft der Exponentialreihe, S.-B. Berlin. Math. Ges., 23 (1924) 50–64.

MEAN VALUE THEOREMS FOR POLYHARMONIC FUNCTIONS

J. H. BRAMBLE AND L. E. PAYNE, University of Maryland

I. Introduction. It is well known that polyharmonic functions satisfy various mean value theorems which may be considered as generalizations of the Gauss mean value theorem for harmonic functions. For example Nicolesco [**3**] gave an expression in terms of certain iterated means and showed that a converse was also true. Cheng [**1**] established a converse for a different mean value expression. Other work on mean value theorems for polyharmonic functions has been carried out by Pizetti [**5**], Picone [**4**], and others (see e.g., [**2**]).

In this paper we derive two rather simple mean value theorems for polyharmonic functions of order p in N dimensions. The expressions involve means over p distinct spheres and seem to be a very natural generalization of the Gauss "peripheral" and "solid" theorems for harmonic functions. A strong converse is given in each case.

We shall consider an N dimensional region R. A function ϕ is called polyharmonic of order p in R if $\phi \in C^{2p}$ and $\Delta^p\phi = 0$ in R where Δ denotes the Laplace operator. For an arbitrary point O of R let S_ρ be the interior of the sphere of radius ρ and center at O. The variable r will be used as the radial variable with respect to O, and the quantity ω_N will denote the surface area of the N dimensional unit sphere.

II. Derivation of the mean value expressions. We start with the following result due to Pizetti [**5**]. Let O be an arbitrary point of R and suppose that $\Delta^p\phi = 0$ in $S_{\rho_p} \subset R$. Then for any $\rho_j \leqq \rho_p$

$$\phi(0) + \sum_{i=2}^{p} \rho_j^{2(i-1)} A_i = \frac{1}{\omega_N} \int_{r=\rho_j} \phi \, d\Omega, \tag{2.1}$$

where the A_i's are independent of ρ_j.

Let the $(p \times p)$ matrix P_{ij} be defined as

$$P_{ij} = \rho_j^{2(i-1)} \tag{2.2}$$

for the p given numbers $0 < \rho_1 < \cdots < \rho_p$, and let P^{ij} be its inverse. Then from (2.1) it follows that

$$\omega_N \phi(0) = \frac{\sum_{j=1}^{p} P^{j1} \int_{r=\rho_j} \phi \, d\Omega}{\sum_{j=1}^{p} P^{j1}}. \tag{2.3}$$

It is easy to see that

$$\sum_{j=1}^{p} P^{j1} = 1. \tag{2.4}$$

Thus

$$\omega_N \phi(0) = \sum_{j=1}^{p} P^{j1} \int_{r=\rho_j} \phi \, d\Omega. \tag{2.5}$$

Because of the form of P_{ij}, and Cramer's rule, we obtain the result in Table A.

$$\omega_N \phi(0) = \frac{\begin{vmatrix} \int_{\rho_1} \phi\, d\Omega & \int_{\rho_2} \phi\, d\Omega & \cdots & \int_{\rho_p} \phi\, d\Omega \\ \rho_1^2 & \rho_2^2 & & \rho_p^2 \\ \rho_1^4 & \vdots & & \vdots \\ \vdots & & & \\ \rho_1^{2(p-1)} & \cdots & & \rho_p^{2(p-1)} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \rho_1^2 & \rho_2^2 & & \rho_p^2 \\ \vdots & \vdots & & \vdots \\ \rho_1^{2(p-1)} & \cdots & & \rho_p^{2(p-1)} \end{vmatrix}} \equiv \frac{D_1}{D_2} \tag{2.6}$$

TABLE A

We can also obtain an expression involving solid means since the identity

$$\phi(0) + \sum_{i=2}^{p} \rho_j^{2(i-1)} B_i = \frac{N}{\rho_j^N \omega_N} \int_{r \leqq \rho_j} \phi \, dV \tag{2.7}$$

can be shown to hold for any function ϕ satisfying $\Delta^p \phi = 0$ in S_{ρ_p}, the B_i's being independent of ρ_j.

In exactly the same way we obtain the result in Table B.

$$\frac{\omega_N}{N} \phi(0) = \frac{\begin{vmatrix} \frac{1}{\rho_1^N} \int_{r \leqq \rho_1} \phi\, dV & \cdots & & \frac{1}{\rho_p^N} \int_{r \leqq \rho_p} \phi\, dV \\ \rho_1^2 & & & \rho_p^2 \\ \vdots & & & \vdots \\ \rho_1^{2(p-1)} & \cdots & & \rho_p^{2(p-1)} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \rho_1^2 & \rho_2^2 & & \rho_p^2 \\ \vdots & \vdots & & \vdots \\ \rho_1^{2(p-1)} & \cdots & & \rho_p^{2(p-1)} \end{vmatrix}} \tag{2.8}$$

TABLE B

III. Converses. Let

$$\Gamma = \begin{cases} \eta K r^{2p+2-N}, & N \text{ odd or } N > 2p+2 \\ \eta K r^{2p+2-N} \log r, & N \text{ even and } N \leqq 2p+2, \end{cases} \tag{3.1}$$

where K is a constant and η is a nonnegative infinitely differentiable function of r which is 1 in $S_{r_{1/2}}$ and zero outside S_{r_1}. We assume that $S_{r_1} \subset R$.

For each N and p there is a constant K such that

$$(3.2) \qquad v(0) = -\int_R \Gamma \Delta^p v \, dV + \int_{R - S_{r_{1/2}}} v \Delta^p \Gamma \, dV,$$

for every sufficiently smooth v in R. With (3.2) it is easy to prove the following.

THEOREM 1. *Let ϕ be a function integrable over all spheres in R and let ϕ satisfy* (2.6) *almost everywhere, for all $0 < \rho_1 < \cdots < \rho_p$, with ρ_p sufficiently small. Then ϕ is equal, almost everywhere, to a function $\bar{\phi}$ which is polyharmonic of order p.*

Proof. Let $\rho_2, \cdots, \rho_p$ be fixed and keep $0 < r < r_1 < \rho_2$. Then, from (2.6),

$$(3.3) \qquad \omega_N \phi(0) = \int_{r_{1/2} < r < r_1} D_2 \Delta^p \Gamma r^{N-1} \, dr = \int_{r_{1/2} < r < r_1} D_1 \Delta^p \Gamma r^{N-1} \, dr.$$

Expanding D_1 and D_2 by means of their first columns, we observe that every term except the first in each case vanishes, because

$$(3.4) \qquad \omega_N \int_{r_{1/2} < r < r_1} \Delta^p \Gamma r^{2i} r^{N-1} \, dr = \int_{r_{1/2} < r < r_1} r^{2i} \Delta^p \Gamma \, dV,$$

$i = 1, \cdots, p-1$ (note that Γ depends only on r). Since $\Delta^p r^{2i} = 0$ and $r^{2i} = 0$ for $r = 0$, $i = 1, \cdots, p-1$, we conclude from (3.2) and (3.4), by setting $v = r^{2i}$, that

$$(3.5) \qquad \int_{r_{1/2} < r < r_1} \Delta^p \Gamma r^{2i} r^{N-1} \, dr = 0, \qquad i = 1, \cdots, p-1.$$

Hence (3.3) reduces to

$$(3.6) \qquad \phi(0) = \int_{r_{1/2} < r < r_1} \phi \Delta^p \Gamma \, dV, \text{ almost everywhere.}$$

Since $\Delta^p \Gamma$ is infinitely differentiable for $r_{1/2} < r < r_1$ and a relation such as (3.6) holds almost everywhere for all sufficiently small r, it follows by standard arguments that there is an infinitely differentiable function $\bar{\phi}$ such that $\phi = \bar{\phi}$ almost everywhere in R. Hence (3.6) holds for $\bar{\phi}$. But from (3.2) we have

$$(3.7) \qquad \int_{r < r_1} \Gamma \Delta^p \bar{\phi} \, dV = \int_R \Gamma \Delta^p \bar{\phi} \, dV = 0.$$

Now Γ is of one sign and r is arbitrarily small, so that $\Delta^p \bar{\phi}(0)$ must be zero. Since O is an arbitrary point it follows that $\Delta^p \bar{\phi} = 0$ in R and the theorem is proved.

THEOREM 2. *Let ϕ be a locally integrable function in R and satisfy* (2.8) *almost everywhere for all $0 < \rho_1 < \cdots < \rho_p$ with ρ_p sufficiently small. Then ϕ is equal almost everywhere to a function $\bar{\phi}$ which is polyharmonic of order p.*

Proof. Multiplying numerator and denominator by ρ_1^N of (2.8), and differentiating with respect to ρ_1, we obtain the result in Table C.

$$
(3.8)\qquad \begin{vmatrix} 1 & \cdots & 1 \\ \frac{(N+2)}{N}\rho_1^2 & \cdots & \rho_p^2 \\ \vdots & & \vdots \\ \frac{[N+2(p-1)]}{N}\rho_1^{2(p-1)} & \cdots & \end{vmatrix} \omega_N\phi(0) = \begin{vmatrix} \int_{r=\rho_1}\phi\, d\Omega & \frac{1}{\rho_2^N}\int_{r\leqq\rho_2}\phi\, dV \cdots & \frac{1}{\rho_p^N}\int_{r\leqq\rho_p}\phi\, dV \\ (N+2)\rho^2 \cdots & & \rho_p^2 \\ \vdots & & \vdots \\ (N+2(p-1))\ \rho_1^{2(p-1)} \cdots & & \end{vmatrix}
$$

TABLE C

Note that (2.8) implies that ϕ is bounded almost everywhere and hence there is a bounded function $\bar{\phi}$ for which (2.8) is satisfied and $\rho=\bar{\phi}$ almost everywhere. Also note that if ρ_1 is small enough then the determinant on the left is not zero, since for $\rho_1=0$ it is a Vandermonde determinant which is different from zero if $0<\rho_2<\cdots<\rho_p$.

Just as before we now obtain (3.6) from (3.8).

The research of the first author was supported in part by the National Science Foundation under Grant NSF GP-2284; that of the second author was supported in part by the National Science Foundation under Grant NSF GP-3.

References

1. M. Cheng, On a theorem of Nicolesco and generalized Laplace operators, Proc. Amer. Math. Soc., 2 (1951) 77–86.

2. Ghermanesco, Sur les moyennes successives des fonctions, Bull. Soc. Math. France, 52 (1934) 245–264.

3. M. Nicolesco, Les fonctions polyharmoniques, Actualités Sci. Ind., 4 (1936).

4. M. Picone, Nuovi indirizzi di ricerca teoria e nel calcolo soluzioni di talune equazioni lineari alle derivate parziali della Fisica-Matematica, Ann. Scuola Norm. Sup. Pisa, 4 (1935) 213–288.

5. P. Pizetti, Sul significato geometrico del secondo parametro differenziale di una funzione sopra una superficie qualunque, Rend. Lincei, 18 (1909) 309–316.

SOLUTIONS OF CERTAIN NONLINEAR DIFFERENTIAL EQUATIONS BY A SEPARATION OF VARIABLES

J. L. TOLLEFSON, University of Idaho, AND F. M. STEIN, Colorado State University

1. Certain nonlinear differential equations can be solved by the method of separation of variables under a suitable transformation of the dependent variable. An important general differential equation which we shall discuss in this paper, and which contains the linear equation ($n=1$), Riccati's equation ($n=2$), and Abel's equation ($n=3$) as special cases, is

$$y'(x) = \sum_{i=0}^{n} f_i(x)y^i(x), \quad (n \text{ a positive integer}). \tag{1}$$

We shall present a transformation which under certain conditions will transform (1) into an equation in which the variables are separable.

2. We assume that f_0 and f_n are not equal to zero for x in the interval under discussion, that they possess first derivatives, and that $f_0 f_n > 0$ if n is even. Upon making the substitution

$$y = (\sqrt[n]{f_0/f_n})v \tag{2}$$

in (1) we obtain

$$\begin{aligned} v' = {} & \sqrt[n]{f_n f_0^{n-1}} + (f_1 - f_0'/nf_0 + f_n'/nf_n)v + f_2(f_0/f_n)^{1/n}v^2 + \cdots \\ & + f_r(f_0/f_n)^{(r-1)/n}v^r + \cdots + (\sqrt[n]{f_n f_0^{n-1}})v^n. \end{aligned} \tag{3}$$

We are led to the following result.

THEOREM. *A differential equation of the form*

$$y'(x) = \sum_{i=0}^{n} f_i(x)y^i(x)$$

with real coefficients and n a positive integer and for which the following conditions hold:

(i) *f_0 and f_n possess first derivatives,*

(ii) *$f_0 f_n > 0$ when n is even, and $f_0 f_n \neq 0$ when n is odd (for x in the interval under discussion),*

(iii) *there exist $n-1$ constants, K_r, such that*

$$K_1 = \frac{f_1 - f_0'/nf_0 + f_n'/nf_n}{\sqrt[n]{f_n f_0^{n-1}}},$$

and

$$K_r = (\sqrt[n]{f_0^{r-n} f_n^{-r}})f_r, \qquad r = 2, 3, \cdots, n-1,$$

can be transformed by the change of variable $y=(\sqrt[n]{f_0/f_n})v$ *to*

$$v' = \sqrt[n]{f_n f_0^{n-1}}(1 + K_1 v + K_2 v^2 + \cdots + K_{n-1} v^{n-1} + v^n),$$

in which the variables are separable.

3. A different, but more restrictive, transformation along the same line has been given by Rao. In [1] he presented the transformation

$$y = vu - f_1/f_2 \tag{4}$$

which reduces Riccati's equation, under certain conditions on the coefficients, to one in which the variables are separable. He generalized (4) in [2] where he presented the transformation

$$y = vu - f_{n-1}/nf_n \tag{5}$$

which, under certain conditions (see (8)), reduces the differential equation (1) to one in which the variables are separable.

For simplicity of notation Rao defined in [2]:

$$\begin{aligned} U(x) &= \sum_{j=0}^{n} (-1)^j n^{-j} f_{n-1}^j f_j f_n^{n-1-j} + (f_n^{n-3}/n)(f_n f'_{n-1} - f_{n-1} f'_n), \\ J(x) &= \Big(v \sum_{j=1}^{n} (-1)^{j-1} j n^{1-j} f_j f_{n-1}^{j-1} f_n^{n-j}\Big) - f_n^{n-1} v', \\ G_r(x) &= \sum_{j=r}^{n} (-1)^{j-r} n^{r-j} \binom{j}{r} f_j f_{n-1}^{j-r} f_n^{n-j-1}. \end{aligned} \tag{6}$$

Upon substituting (5) in (1) he obtained for $n \geqq 4$,

$$f_n^{n-1} v u' = U(x) + uJ(x) + \sum_{r=2}^{n-2} f_n v^r G_r(x) u^r + f_n v^n u^n. \tag{7}$$

Under the assumption that $U \neq 0$, he then defined v by $f_n^n v^n = U$ on the interval where $U>0$. If $n-2$ constants (in [2] Rao referred to $n-1$ constants instead of $n-2$) C_r exist, such that

$$\begin{aligned} (1/n) f_n^{n-2} U' - \Big(f_n^{n-3} f_n' + \sum_{j=1}^{n} (-1)^{j-1} j n^{1-j} f_j f_{n-1}^{j-1} f_n^{n-j-1}\Big) U &= C_1 U^{2-1/n}, \\ G_r U^{r/n-1} = C_r, \qquad r = 2, 3, \cdots, n-2; \end{aligned} \tag{8}$$

then Rao obtained from (7) the following equation in which the variables are separable:

$$u' = (f_n^{2-n} U^{1-1/n})(1 - C_1 u + C_2 u^2 + \cdots + C_{n-2} u^{n-2} + u^n). \tag{9}$$

Rao noted that (7) is also valid for $n=2, 3$, if the third term on its right side is

taken equal to zero, and for $n=1$ if the last two terms on its right side are taken equal to zero.

For the transformation (5) to apply, Rao required that f_{n-1} and f_n be twice differentiable, that all f_i ($i=0, 1, \cdots, n-2$) be differentiable, and that $f_n>0$.

4. Rao's transformation (5) is more restrictive than the transformation (2) which we present as far as conditions for differentiability are concerned. Observe that (2) requires only that f_0 and f_n possess first derivatives, while (5) requires all f_i's to be once differentiable and f_{n-1} and f_n to be twice differentiable. See section 5 for a case in which Rao's transformation fails to apply because of this restriction. Rao's condition that $f_n>0$ corresponds to the condition that $f_0 f_n>0$ if n is even, and that $f_0 f_n \neq 0$ if n is odd.

It should be noted that in [**3**] Allen and Stein presented a transformation that reduces Riccati's equation to one with separable variables under certain conditions (less restrictive than Rao's for this case; see [**1**]). The transformation by Allen and Stein,

$$y = \sqrt{f_0/f_2}\, v, \tag{10}$$

is a special case of (2), when $n=2$.

5. An example of the application of (2) is the following equation of Abel's type in which the coefficients satisfy the conditions of the theorem,

$$y' = \exp(3/2\, x^2) + [\exp(x^2 + 2/3\, x^{3/2}) - \sqrt{x} + x]y + \exp(2x^{3/2})y^3. \tag{11}$$

The substitution (2), $y=[\exp(x^2/2-2/3x^{3/2})]v$, transforms equation (13) to

$$v' = \exp(x^2 + 2/3\, x^{3/2})(1 + v + v^3). \tag{12}$$

Notice that Rao's transformation (5) does not apply to (11) on an interval containing the origin because $f_3(x)$ is not twice differentiable and $f_2(x)$ is not once differentiable on such an interval. Even in an interval not containing $x=0$, the work involved in Rao's transformation is quite involved.

Prepared by Mr. Tollefson, under the direction of Professor Stein in an NSF Undergraduate Science Education Program.

References

1. P. R. P. Rao, The Riccati differential equation, this MONTHLY, 69 (1962) 995.

2. ———, A differential equation, this MONTHLY, 71 (1964) 530–533.

3. James L. Allen and F. Max Stein, On solutions of certain Riccati differential equations, to be published in this MONTHLY.

4. George M. Murphy, Ordinary differential equations and their solutions, Van Nostrand, Princeton, 1960.

MEASURE AND DENSITY OF SEQUENCES

WILLIAM HINTZMAN, University of Wisconsin

Let I denote the set of positive integers and $\mathcal{P}(I)$, the collection of all subsets of I. If $S \in \mathcal{P}(I)$, let $S(n)$ equal the number of elements in $\{x \in S : x \leqq n\}$. Let

$$\mu^*(S) = \limsup_{n \to \infty} \frac{S(n)}{n} \cdot$$

$\mu^*(S)$ may be called the upper density of S in I. Then a finite set has upper density 0 while an arithmetical progression $\{an+b\}$ has upper density $1/a$.

If A, $B \in \mathcal{P}(I)$, then

$$\frac{(A \cup B)(n)}{n} \leqq \frac{A(n)}{n} + \frac{B(n)}{n} \cdot$$

Hence, $\mu^*(A \cup B) \leqq \mu^*(A) + \mu^*(B)$. Also, if $A \subset B$, then $A(n)/n \leqq B(n)/n$ and $\mu^*(A) \leqq \mu^*(B)$. Therefore, μ^* is a finitely subadditive outer measure on $\mathcal{P}(I)$. Let $\mathfrak{M}$ be the collection of all sets $S \in \mathcal{P}(I)$ such that for all $X \in \mathcal{P}(I)$, $\mu^*(X) = \mu^*(X \cap S) + \mu^*(X \cap S')$. From classical measure theory we know that $\mathfrak{M}$ is an algebra of sets and that μ^* is a finitely additive measure on $\mathfrak{M}$.

An inner measure μ_* on $\mathcal{P}(I)$ is obtained by setting $\mu_*(S)$ equal to

$$\mu^*(I) - \mu^*(S') = \liminf \frac{S(n)}{n} \cdot$$

Let $\mathfrak{D} = \{S \in \mathcal{P}(I) : \mu^*(S) = \mu_*(S)\}$. Then $S \in \mathfrak{D}$ if and only if S has density, i.e., if and only if $D(S) = \lim S(n)/n = \mu^*(S) = \mu_*(S)$ [**1**]. It is clear that $\mathfrak{M} \subset \mathfrak{D}$ for if $S \in \mathfrak{M}$, $\mu^*(I) = \mu^*(S) + \mu^*(S')$.

THEOREM 1. *If* $A \in \mathfrak{D}$ *and* $D(A) = 1$ *or* $D(A) = 0$, *then* $A \in \mathfrak{M}$.

Now $D(A) = 1$ if and only if $D(A') = 0$ for if $A \in \mathfrak{D}$, $1 = D(A) + D(A')$. Suppose $D(A) = 1$ and $X \in \mathcal{P}(I)$. Since $(X \cap A) \subset X$ and $(X \cap A') \subset A'$, $\mu^*(X \cap A) \leqq \mu^*(X)$ and $0 \leqq \mu^*(X \cap A') \leqq \mu^*(A') = 0$. Therefore, $\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap A')$ and A, $A' \in \mathfrak{M}$.

THEOREM 2. $\mathfrak{M} \neq \mathfrak{D}$, *for the only sets in* $\mathfrak{M}$ *are those of density* 0 *or* 1.

Suppose that $S \in \mathfrak{M}$ and $D(S) = a$, $D(S') = 1 - a$, $a \neq 1, 0$. Let

$$X = \{n : (4k)^{4k} \leqq n \leqq (4k+1)^{4k+1}, \qquad (k = 1, 2, \cdots)\},$$
$$Y = \{n : (4k+2)^{4k+2} \leqq n \leqq (4k+3)^{4k+3}, \qquad (k = 0, 1, 2, \cdots)\}.$$

Then

$$\mu^*(X) = \mu^*(Y) = \mu^*(X \cup Y) = 1, \qquad X \cap Y = \phi.$$

Since $S \in \mathfrak{M}$, $1 = \mu^*(X) = \mu^*(X \cap S) + \mu^*(X \cap S')$ and we have

$$\mu^*(X \cap S) = a, \qquad \mu^*(X \cap S') = 1 - a.$$
$$\mu^*(Y \cap S) = a, \qquad \mu^*(Y \cap S') = 1 - a.$$

Let $A = (X \cap S) \cup (Y \cap S')$; then

$$\mu^*(A \cap S) + \mu^*(A \cap S') = \mu^*(X \cap S) + \mu^*(Y \cap S') = a + 1 - a = 1 > \mu^*(A),$$

since $\mu^*(A) = \max\,[a, 1-a]$. This contradiction proves the assertion.

References

1. Ivan Niven, The asymptotic density of sequences, Bull. Amer. Math. Soc., 57 (1951) 420–434.

2. R. C. Buck, The measure-theoretic approach to density, Amer. J. Math., 68 (1946) 560–580.

QUASI-INVERSES OF SEQUENCES

R. G. BUSCHMAN, State University of New York at Buffalo

In the convolution ring discussed by Louis Brand [1] there is a close relationship between the difference equations for a sequence, U, generated from the numbers u_0, u_1 by

$$u_{n+2} = au_{n+1} + bu_n, \qquad n \geqq 0,$$

and its quasi-inverse, V. Since the ring has an identity I, the quasi-inverse V of U, can be defined as the solution in the convolution ring of $(I-U)(I-V) = I$ when $u_0 \neq 1$.

From the generating function for U, given by

$$g(U; s) = [u_0 + (u_1 - au_0)s]/(1 - as - bs^2),$$

we can obtain the generating function for V by successively computing $g(I-U; s)$, $g(I-V; s)$, and then $g(V; s)$. This yields

$$g(V; s) = \frac{[-u_0(1 - u_0)] - [(u_1 - au_0)/(1 - u_0)]s}{1 - [(a + u_1 - au_0)/(1 - u_0)]s - [b/(1 - u_0)]s^2}$$

so that the difference equation is of the same form, $v_{n+2} = Av_{n+1} + Bv_n$, $n \geqq 0$, where we have

$$v_0 = -u_0/(1 - u_0),\ v_1 = -u_1/(1 - u_0)^2,\ A = a + u_1/(1 - u_0),\ B = b/(1 - u_0).$$

Two special cases of interest follow as examples. If $u_0 = 0$, the difference equation for V shows an even stronger similarity:

$$0, \quad -u_1, \quad v_{n+2} = (a + u_1)v_{n+1} + bv_n, \qquad n \geqq 0.$$

Here the coefficient of the lowest order term is the same as for U and the other coefficient is a simple translation of that of U. In particular, for the Fibonacci sequence, $F=\{0, 1, 1, 2, \cdots\}$, the quasi-inverse becomes

$$0, \quad -1, \quad v_{n+2} = 2v_{n+1} + v_n, \qquad n \geqq 0.$$

If $u_0=2$, the expression for V becomes

$$2, \quad -u_1, \quad v_{n+2} = (a - u_1)v_{n+1} - bv_n, \qquad n \geqq 0;$$

so that the quasi-inverse for the Lucas sequence, $L=\{2, 1, 3, 4, \cdots\}$, is given by

$$2, \quad -1, \quad v_{n+2} = - v_n, \qquad n \geqq 0.$$

Similar situations occur for higher order difference equations with a wider variety of interesting special cases thus occurring. Here again if the initial values for U satisfy certain conditions, then one or more of the coefficients of the difference equation for V are the same as those of U; the others are simple translations.

The inverse sequence, U^{-1}, which exists for $u_0 \neq 0$, does not seem to possess interesting properties of this type.

Reference

1. Louis Brand, A division algebra for sequences and its associated operational calculus, this MONTHLY, 71 (1964) 719–728.

COUNTEREXAMPLES IN HAAR MEASURE

S. K. BERBERIAN, The University of Iowa

The writer on abstract harmonic analysis rightly regards integration theory and Haar measure as technical preliminaries to the main business at hand, and he is entitled, perhaps even obligated, to treat such matters summarily. The danger is that these technicalities underlie the theory of convolution, which is very delicate, and the result is that mistakes are not only common, they are almost traditional. The overall mathematical effect of such errors is usually negligible, but they can be disturbing to the reader. My aim in this expository article is to arm the reader against two of the most common errors, so that when he meets them he may take them in his stride.

To be more specific, the purpose of the article is to publicize the following two propositions:

PROPOSITION A. *The product of two regular Borel measures may fail to be a regular Borel measure.*

PROPOSITION B. *Haar measure may fail to be simultaneously inner and outer regular.*

These propositions have been known for some time, but I have not traced them to their origins, and the purpose of the bibliography at the end of the article is merely to supply factual references. I learned them mainly from the book of P. R. Halmos, and from conversations with Roy A. Johnson and George Burke.

1. Our terminology is drawn from [**1**] and [**2**]. Let X be a locally compact (tacitly Hausdorff) topological space. The class of *Baire sets* in X is the σ-ring generated by the compact G_δ's, and is denoted $\mathcal{B}_0(X)$. The class of *Borel sets* in X is the σ-ring generated by the compact sets, and is denoted $\mathcal{B}(X)$. The class of *weakly Borel sets* in X is the σ-ring generated by the closed sets; it is a σ-algebra, and is denoted $\mathcal{B}_w(X)$. The Borel sets are precisely the σ-bounded weakly Borel sets [**1**, p. 181].

When X is metrizable, $\mathcal{B}_0(X)=\mathcal{B}(X)$; but there exist non-metrizable compact spaces X for which the equality holds [**3**]. Clearly $\mathcal{B}(X)=\mathcal{B}_w(X)$ if and only if X is σ-compact.

2. If $\mathcal{S}$ and $\mathcal{T}$ are any two σ-rings, then $\mathcal{S}\times\mathcal{T}$ denotes the σ-ring generated by all "rectangles" of the form $E\times F$, with "sides" in $\mathcal{S}$ and $\mathcal{T}$, respectively [**2**, p. 140]. For any pair of locally compact spaces X and Y, the following relations hold:

$$\mathcal{B}_0(X) \times \mathcal{B}_0(Y) = \mathcal{B}_0(X \times Y), \tag{1}$$

$$\mathcal{B}(X) \times \mathcal{B}(Y) \subset \mathcal{B}(X \times Y), \tag{2}$$

$$\mathcal{B}_w(X) \times \mathcal{B}_w(Y) \subset \mathcal{B}_w(X \times Y). \tag{3}$$

The proofs of (2) and (3) are no deeper than the assertion that a rectangle with compact [resp. closed] sides is compact [resp. closed] (cf. [**1**, p. 118]). The proof of (1) is more delicate [**2**, p. 222]. Informally, we may say that "Baire sets multiply." In general, Borel sets and weakly Borel sets do not; in fact, there exist compact topological groups G for which the inclusion

$$\mathcal{B}(G) \times \mathcal{B}(G) \subset \mathcal{B}(G \times G)$$

is proper (cf. [**2**, p. 261 ff.] and [**3**]). Incidentally, for a locally compact group G, the condition $\mathcal{B}_0(G)=\mathcal{B}(G)$ is equivalent to metrizability (cf. [**2**, p. 221]).

3. Suppose that X is any topological space, $\mathcal{S}$ is a σ-ring of subsets of X, and τ is a measure defined on $\mathcal{S}$. We say that τ is *inner regular* if

$$\tau(E) = \text{l.u.b.}\ \{\tau(C)\colon C \subset E,\ C \in \mathcal{S},\ C \text{ compact}\} \tag{4}$$

for each E in $\mathcal{S}$; *outer regular* if

$$\tau(E) = \text{g.l.b.}\ \{\tau(U)\colon\ E \subset U,\ U \in \mathcal{S},\ U \text{ open}\} \tag{5}$$

for each E in $\mathcal{S}$; and *biregular* if it is both inner and outer regular.

4. Let X be a locally compact space. A *Baire measure* on X is a measure ν defined on $\mathfrak{B}_0(X)$ such that $\nu(D) < \infty$ for all compact G_δ's D. A *Borel measure* on X is a measure μ defined on $\mathfrak{B}(X)$ such that $\mu(C) < \infty$ for all compact sets C. A *weakly Borel measure* on X is a measure ρ defined on $\mathfrak{B}_w(X)$ such that $\rho(C) < \infty$ for all compact sets C. It is known that (a) every Baire measure is biregular, (b) a Borel measure is inner regular if and only if it is outer regular, and (c) it can happen that a weakly Borel measure is either inner regular or outer regular, but not both. (Cf. [**1**, p. 194], [**2**, p. 228], and Sections 7 and 8 below.)

A Borel measure is said to be *regular* if it is inner regular in the sense of (4); thus every regular Borel measure is biregular. A weakly Borel measure ρ is said to be *regular* if it is inner regular in the sense of (4); thus the restriction of ρ to the class of Borel sets is a regular Borel measure, but ρ itself need not be biregular.

If ν is a Baire measure, there exists one and only one regular Borel measure μ which extends ν [**2**, p. 239]. If μ is a regular Borel measure, there exists one and only one regular weakly Borel measure ρ which extends μ, namely

$$\rho(A) = \text{l.u.b.}\ \{\mu(C)\colon\ C \subset A,\ C \text{ compact}\}$$

for each weakly Borel set A [**1**, p. 203 ff.].

5. Let G be a locally compact topological group. We discuss (left invariant) Haar measure on three levels: Baire, Borel, and weakly Borel.

Haar measure on the Baire sets of G is a Baire measure ν, not identically zero, such that $\nu(xF) = \nu(F)$ for all Baire sets F, and all x in G. Such a measure ν exists, and is unique up to a factor of proportionality [**1**, p. 261]. It follows that $\nu(V) > 0$ for every nonempty open Baire set V.

Haar measure on the Borel sets of G is a regular Borel measure μ, not identically zero, such that $\mu(xE) = \mu(E)$ for all Borel sets E, and all x in G. Such a measure μ exists, and is unique up to proportionality [**1**, p. 260]. It follows that $\mu(U) > 0$ for every nonempty open Borel set U. It turns out that the assumption of regularity is redundant, but this fact is far from obvious [**1**, p. 299].

Haar measure on the weakly Borel sets of G is a regular weakly Borel measure ρ, not identically zero, such that $\rho(xA) = \rho(A)$ for all weakly Borel sets A, and all x in G. Such a measure ρ exists, and is unique up to proportionality [**1**, p. 263 ff.]. It follows that $\rho(W) > 0$ for every nonempty open set W. The assumption of regularity is in general not redundant; an example is given in Section 8.

6. *Proof of Proposition A*. Let X be any locally compact space for which the inclusion

$$\mathfrak{B}(X) \times \mathfrak{B}(X) \subset \mathfrak{B}(X \times X)$$

is proper, and let μ be any regular Borel measure on X. Then the product measure $\mu \times \mu$, as defined in [**2**], cannot be a Borel measure, simply because its domain is too small. In view of the remarks in Section 2, it does not help to assume that X is a compact group and that μ is Haar measure.

7. *Proof of Proposition B.* More precisely, we prove:

THEOREM 1. *Let ρ be Haar measure on the weakly Borel sets of a locally compact group G. In order that ρ be biregular, it is necessary and sufficient that G be either discrete or σ-compact.*

Proof of necessity. Assuming that G is neither discrete nor σ-compact, let us show that ρ cannot be biregular. Choose a compact neighborhood of the identity, and let H be the subgroup generated by this neighborhood. Then H is open, closed, and σ-compact [**2**, p. 251]. Since G is not σ-compact, there must be uncountably many left cosets. Let S be a subset of G which intersects each left coset in exactly one point (Axiom of Choice). We shall see that S is a closed set, its compact subsets have measure zero, and its open supersets have infinite measure; this will preclude the biregularity of ρ.

Observe that if C is any compact set in G, then $C \cap S$ is finite; this follows from the fact that the left cosets form an open covering of G. In particular, every compact subset of S is finite.

Suppose that x is a point of G not in S. Choose a compact neighborhood U of x. Since $U \cap S$ is a finite set excluding x, and since G is a Hausdorff space, we may find a neighborhood V of x such that $V \cap (U \cap S)$ is empty. Then $V \cap U$ is a neighborhood of x which is disjoint from S, and so x is not adherent to S. Thus S is closed, and is therefore a weakly Borel set.

If C is any compact subset of S, then C is finite, and it follows that $\rho(C) = 0$ (recall that G is not discrete; cf. [**2**, p. 268 ff., Exercise 3]).

On the other hand, if W is any open set containing S, its intersection with each left coset is nonempty and open; thus W is partitioned into an uncountable family of sets of positive measure, and therefore has infinite measure.

Proof of sufficiency. If G is σ-compact, then $\mathcal{B}_w(G) = \mathcal{B}(G)$; thus ρ is a regular Borel measure, and is therefore biregular.

Finally, if G is discrete, then all subsets are open and so outer regularity is trivial in this case.

8. Suppose that G is a locally compact group which is neither discrete nor σ-compact, and let μ be Haar measure on the Borel sets of G. One can extend μ to a left invariant weakly Borel measure by assigning infinite measure to the weakly Borel sets which are not Borel sets. The resulting measure is outer regular, but not inner regular.

9. In view of Theorem 1, there is in general no remedy for Proposition B; one either learns to live with lopsidedly regular Haar measure on weakly Borel sets, or else one retreats to a smaller class of measurable sets (e.g. Borel sets or Baire sets).

For Proposition A, there are at least two remedies. The simplest is to stick to Baire measures, and lean on the relation (1); this is drastic, but often practicable (cf. [**1**, Chapter 9]). Another is as follows (cf. [**3**]):

THEOREM 2. *If μ_1 and μ_2 are regular Borel measures on the locally compact spaces X_1 and X_2, respectively, then there exists one and only one regular Borel measure μ on $X_1 \times X_2$ which extends $\mu_1 \times \mu_2$.*

It is instructive to prove a slightly more general result, which is also useful in the theory of spectral measures:

THEOREM 3. *Let X_1 and X_2 be locally compact spaces, and suppose that τ is a measure on the σ-ring $\mathcal{B}(X_1) \times \mathcal{B}(X_2)$ such that* (i) *for each compact set C_1 in X_1, the correspondence*

$$E_2 \to \tau(C_1 \times E_2) \qquad (E_2 \in \mathcal{B}(X_2))$$

is a regular Borel measure on X_2, and (ii) *for each compact set C_2 in X_2, the correspondence*

$$E_1 \to \tau(E_1 \times C_2) \qquad (E_1 \in \mathcal{B}(X_1))$$

is a regular Borel measure on X_1. Then τ may be extended to one and only one regular Borel measure μ on $X_1 \times X_2$.

Proof. The uniqueness of μ follows from the fact that the domain of definition of τ includes the Baire sets of $X_1 \times X_2$ (cf. formula (1)).

Let $\mathcal{R}_1$ be the ring generated by the compact sets in X_1. Every set in $\mathcal{R}_1$ is a finite disjoint union of "proper differences" $C_1 - C_1^*$, where C_1 and C_1^* are compact sets such that $C_1^* \subset C_1$ [**2**, p. 223]. Similarly, let $\mathcal{R}_2$ be the ring generated by the compact sets in X_2.

Let $\mathcal{R}$ be the ring generated by the class of all rectangles $E_1 \times E_2$ with sides in $\mathcal{R}_1$ and $\mathcal{R}_2$, respectively. Each set in $\mathcal{R}$ can be written as a finite disjoint union of sets of the form

$$(C_1 - C_1^*) \times (C_2 - C_2^*),$$

where both of the indicated differences are proper (cf. [**2**, p. 139]). Such a set can be written in the form

$$(C_1 \times C_2 - C_1 \times C_2^*) - (C_1^* \times C_2 - C_1^* \times C_2^*), \tag{6}$$

where each of the indicated differences is proper.

It is implicit in the assumption (i) that $\tau(C_1 \times C_2) < \infty$ for all rectangles with compact sides, and it follows that the restriction of τ to the class of Baire sets of $X_1 \times X_2$ is a Baire measure ν. Let μ be the unique regular Borel extension of ν. Our problem is to show that

$$\mu(E) = \tau(E) \tag{*}$$

for all sets E in $\mathcal{B}(X_1) \times \mathcal{B}(X_2)$. It is sufficient to verify (*) for rectangles $E = C_1 \times C_2$ with compact sides; for then it will follow from (6) that (*) holds for all sets in $\mathcal{R}$ [**2**, p. 37], and therefore for all sets in the σ-ring generated by $\mathcal{R}$ [**2**, p. 54], in other words for all sets in $\mathcal{B}(X_1) \times \mathcal{B}(X_2)$ [**1**, p. 118].

We are thus reduced to verifying (*) for a given compact rectangle $E=C_1\times C_2$. By (i) we may choose a compact G_δ D_2 in X_2 such that $C_2\subset D_2$ and

$$\tau(C_1\times C_2)=\tau(C_1\times D_2)$$

(cf. [**1**, p. 188]). In turn, by (ii) there exists a compact G_δ D_1 in X_1 such that $C_1\subset D_1$ and

$$\tau(C_1\times D_2)=\tau(D_1\times D_2).$$

Thus $C_1\times C_2\subset D_1\times D_2$, and

$$\tau(C_1\times C_2)=\tau(D_1\times D_2). \tag{7}$$

On the other hand, by the regularity of μ, there exists a compact G_δ D in $X_1\times X_2$ such that $C_1\times C_2\subset D$ and

$$\mu(C_1\times C_2)=\mu(D). \tag{8}$$

Let $D^*=D\cap(D_1\times D_2)$. Then D^* is a Baire set (in fact, a compact G_δ), and $C_1\times C_2\subset D^*$. Since μ and τ agree on Baire sets, we have, citing (7),

$$\mu(C_1\times C_2)\leqq\mu(D^*)=\tau(D^*)\leqq\tau(D_1\times D_2)=\tau(C_1\times C_2);$$

similarly, citing (8), we have $\tau(C_1\times C_2)\leqq\tau(D^*)=\mu(D^*)\leqq\mu(D)=\mu(C_1\times C_2)$. Thus $\mu(C_1\times C_2)=\tau(C_1\times C_2)$, and Theorem 3 is completely proved.

To derive Theorem 2, we simply apply Theorem 3 to the product measure $\tau=\mu_1\times\mu_2$; conditions (i) and (ii) are verified using the fact that

$$\tau(E_1\times E_2)=\mu_1(E_1)\mu_2(E_2)$$

for all rectangles with Borel sides (cf. [**2**, p. 144], and [**1**, p. 32, Lemma 2]).

Finally, we remark that Theorem 2 is by no means the end of the story; the task of successfully involving μ_1, μ_2, and μ in theorems of Fubini-Tonelli type is far from trivial [**3**].

References

1. S. K. Berberian, Measure and integration, Macmillan, New York, 1965.

2. P. R. Halmos, Measure theory, Van Nostrand, New York, 1950.

3. R. A. Johnson, On product measures and Fubini's theorem in locally compact spaces, University of Iowa Dissertation, Iowa City, 1964.

MAXIMAL AND MINIMAL SUBADDITIVE EXTENSIONS

MARSHALL BARTON AND RICHARD LAATSCH, Miami University, Oxford, Ohio

1. Introduction. A function f with domain D, a subset of the set R of all real numbers, is called subadditive on D if

$$f(x + y) \leqq f(x) + f(y)$$

whenever x, y, and $x+y$ belong to D. This paper is concerned with two extensions of subadditive functions, namely the maximal subadditive extension from an interval $[0, a](a>0)$ to the set $E=[0, \infty)$ and the minimal subadditive extension from E to R. The maximal extension was first defined by Bruckner [1] and was discussed also in [2]. The minimal extension is first defined below.

2. The maximal extension. Let f be subadditive on the interval $[0, a]$. Then the *maximal subadditive extension* of f to $E=[0, \infty)$ is the function Sf defined at each $x \in E$ by $Sf(x)=\inf \sum f(x_i)$, where the infimum is taken over all finite collections $\{x_1, x_2, \cdots, x_n\}$ such that $0 \leqq x_i \leqq a$ $(i=1, 2, \cdots, n)$ and $\sum x_i = x$. The collection $\{x_1, \cdots, x_n\}$ is called an *a-partition* of x. It can be readily verified that Sf is subadditive on E and, if g is any subadditive extension of f to E, that $g \leqq Sf$. Also, an a-partition of x need not be an arbitrarily large set, since we can replace x_i and x_j by x_i+x_j whenever $x_i+x_j \leqq a$ and obtain an approximation to Sf at least as good as the original.

Considerable attention was given in [2] to Sf when $Sf(na+x)=nf(a)+f(x)$ for all $x \in (0, a]$ and all positive integers n (i.e., when Sf "repeats" the behavior of f). Such behavior occurs, for example, when f is concave on $[0, a]$. Let us note in this regard that, if F is subadditive on E and, for some $a>0$, $F(na+x) = nF(a)+F(x)$ for all $x \in (0, a]$ and all positive integers n, then $F=Sf$, where f is the restriction of F to $[0, a]$.

The principal purpose of this section is to establish a necessary and sufficient condition that f have a repeating maximal extension. Geometrically, the condition is as follows: Translate the graph of f on $(0, a]$ by placing the point $(a, f(a))$ at the point $(x, f(x))$. If for no x in $(0, a]$ the translated graph enters the region under the original graph, the extension will repeat, and conversely.

THEOREM 1. *Let f be subadditive on $[0, a]$. Then $Sf(na+x)=nf(a)+f(x)$ for all $x \in (0, a]$ and all positive integers n if, and only if, for all $y \in (0, a]$,*

$$f(y) \leqq f(a + y - u) - f(a) + f(u) \tag{1}$$

for all u satisfying $y \leqq u \leqq a$.

Proof. It will be sufficient to consider only the case $n=1$ since one repetition implies continued repetition of the behavior of f. (This is Theorem 9 in [2].)

Let $Sf(a+x)=f(a)+f(x)$ for all $x \in (0, a]$. Then, if $0 \leqq y \leqq u \leqq a$,

$$f(a + y - u) \geqq Sf(a + y) - f(u) = f(a) + f(y) - f(u).$$

Conversely, if condition (1) is satisfied, let $x \in (0, a]$, $\epsilon > 0$, and $\{x_1, \cdots, x_n\}$ be an a-partition of $x+a$ such that $Sf(a+x)+\epsilon > \sum f(x_i)$. If $x \leqq x_m$ for some m, then

$$\sum f(x_i) \geqq f(x_m) + f(a + x - x_m) \geqq f(a) + f(x)$$

by condition (1) with $u = x_m$ and $y = x$. If $x > x_i$ for all i, then we can choose x_1 and x_2 from the a-partition such that $0 \leqq x - x_1 \leqq x_2 \leqq a$. Then

$$\sum f(x_i) \geqq f(x_1) + f(x_2) + f(a + x - x_1 - x_2) \geqq f(a) + f(x)$$

by condition (1) with $u = x_2$ and $y = x - x_1$, since also $f(x - x_1) \geqq f(x) - f(x_1)$. In either case, $Sf(a+x)+\epsilon > f(a)+f(x)$ for all $\epsilon > 0$. Since $Sf(a+x) \leqq f(a)+f(x)$ by definition, the proof is complete.

3. The minimal extension. Let f be a subadditive function defined on $E = [0, \infty)$. Let Mf be the extension of f to R defined on $(-\infty, 0)$ by

$$Mf(x) = \sup \{f(u) - f(v) : v > u \geqq 0 \text{ and } u - v = x\}.$$

Then Mf is the *minimal subadditive extension* of f to R. Again it can be shown that Mf is subadditive on R and $Mf \leqq F$ for every subadditive extension F of f to R. The only problem in the proof of this concerns subadditivity when $x+y<0$ in which case we assume that $y<0$, choose $0 \leqq u < v$ such that $u - v = x+y$ and $Mf(x+y) < f(u) - f(v) + \epsilon$, find $w \geqq 0$ such that $u - w = y$, and show that $Mf(x+y) - Mf(y) - \epsilon < Mf(x)$, using the fact that $w - v = x$.

Some immediate consequences of the definition of Mf are that Mf is nondecreasing on R if f is nondecreasing on E, that Mf is nonincreasing on $(-\infty, 0)$ if f is nonincreasing on E (but there may be an interruption at 0), and that, if f is continuous on E, then Mf is continuous on $(-\infty, 0)$ and $Mf(x) \to 0$ as $x \to 0-$. Also, if $x<0$ and $c>0$, then $M(cf(x)) = c(Mf)(x)$ and $M(c+f)(x) = Mf(x)$.

It may be the case that $g \neq c+f$ on E and $Mg = Mf$ on $(-\infty, 0)$. For example, if $f(2n) = f(2n+1) = n+1$ and $g(2n)+1 = g(2n+1) = n+3$ and if f and g are interpolated polygonally on E, then $Mf = Mg$ on $(-\infty, 0)$. (Note also that $M(f+g) \neq Mf + Mg$.)

The next theorem considers boundedness of the graph of Mf by straight lines. (Compare Theorem 10 of [2].)

THEOREM 2. *Let f be subadditive on E. If $m = \inf\{f(x)/x : x > 0\}$ exists in R, then $Mf(x) \geqq mx$ for all $x \in R$. If, additionally, $b = \sup\{f(x) - mx : x \geqq 0\}$ exists in R, then $Mf(x) \leqq mx + b$ for all $x \in R$.*

Proof. Let $x<0$ and $\epsilon>0$. If m exists, choose $v>0$ such that $f(v) < mv + \epsilon$. Since $f(nv) \leqq nf(v)$ for every positive integer n, v can be taken arbitrarily large. Let $u = v + x$. Then

$$Mf(x) \geqq f(u) - f(v) > mu - (mv + \epsilon) = mx - \epsilon.$$

If b also exists, choose $v > u \geqq 0$ such that $u - v = x$ and $Mf(x) < f(u) - f(v) + \epsilon$. Then $f(u) - f(v) \leqq mu + b - mv = mx + b$; so $Mf(x) \leqq mx + b$.

THEOREM 3. *Let $\{f_n\}$ be a sequence of subadditive functions on E converging pointwise on E to f. Then $Mf = \lim Mf_n$.*

We know [**3**] that f is subadditive, and the rest of the proof is relatively straightforward.

From the same reference we have that a concave function f is subadditive on E if, and only if, $f(0) \geqq 0$. For such functions, if the m of Theorem 2 exists, then $Mf(x) = mx$ for every $x < 0$, and $M(f+g) = Mf + Mg$ for such functions, although this is not generally the case.

Finally, to relate the discussion of Mf to our earlier discussion of maximal extensions (especially Theorem 1), let us note that if f repeats on E, then Mf repeats on R. Precisely: Let f be subadditive on $[0, \infty)$ with $f(0) = 0$ and let $a > 0$ be such that $f(na + x) = nf(a) + f(x)$ for all $x \in (0, a]$ and all positive integers n. Then $Mf(na + x) = nf(a) + f(x)$ for all $x \in (0, a]$ and all integers n. (If $f(0) \neq 0$, the conclusion fails only at 0.)

References

1. Andrew Bruckner, Minimal superadditive extensions of superadditive functions, Pacific J. Math., 10 (1960) 1155–1162.

2. Richard G. Laatsch, Extensions of subadditive functions, Pacific J. Math., 14 (1964) 209–215.

3. Einar Hille and R. S. Phillips, Functional Analysis and Semigroups, Amer. Math. Soc. Colloq. Publ., 31 (1957) Chapter 7.

COMPACTNESS IN THE SPACE OF QUASI-CONTINUOUS FUNCTIONS

T. H. HILDEBRANDT, University of Michigan

It is well known that if C is the class of continuous functions on $[a, b]$ with l.u.b. norm, then a subset F of C is compact in the Weierstrass-Bolzano sense (every infinite subset contains a sequence converging uniformly) if and only if the functions f of F are uniformly bounded and equicontinuous at all points of $[a, b]$. Only a slight change is needed to obtain a compactness condition in the space of quasi-continuous functions, those for which $f(x+0)$ and $f(x-0)$ exist for all x of $[a, b]$. We have:

THEOREM. *Necessary and sufficient conditions that a subset F of the class Q of quasi-continuous functions on $[a, b]$ be WB compact are that*: (a) *the functions f of F be uniformly bounded and* (b) *for every x_0 with*

$$a < x_0 \leqq b: \lim_{x\to x_0-0} f(x) = f(x_0 - 0)$$

and every x_0 with $a \leqq x_0 < b$: $\lim_{x\to x_0+0} f(x) = f(x_0+0)$, each uniformly for f in F.

We shall call a property of type (b) equiconvergence.

Necessity. The necessity of the boundedness condition follows in the usual way. For the condition (b) of equiconvergent type, we assume an x_0 where the righthanded condition does not hold. Then there exists an $e>0$, such that for every n, there exists f_n in F and x_n such that $0<x_n-x_0<1/n$ and

$$|f_n(x_n) - f_n(x_0 + 0)| > e,$$

that is no subsequence of $f_n(x)$ has the rightconvergence property at x_0. Since F is compact, there exists a subsequence $f_{n_m}(x)$ converging uniformly on $[a, b]$ and consequently, by the iterated limits theorem [**1**, p. 14], having the right equiconvergence property at x_0. This contradicts the assumption on $f_n(x)$.

Sufficiency. Let $f_n(x)$ be an infinite sequence in F and $\{r_m\}$ a denumerable dense set on $[a, b]$ including a and b. Then since by condition (a) for every m, the sequence $f_n(r_m)$ is bounded, there exists a subsequence such that $\lim_k f_{n_k}(r_m)$ exists for all m [**1**, pp. 44–5]. For convenience we shall designate this sequence f_k. If x_0 is any point of $[a, b]$ then by condition (b)

$$\lim_{r_m\to x_0-0} f_k(r_m) = f_k(x_0 - 0)$$

uniformly in k, where $r_m<x_0$. Consequently by the iterated limits theorem

$$\lim_{r_m\to x_0-0} \lim_k f_k(r_m) = \lim_k f_k(x_0 - 0).$$

Similarly

$$\lim_{r_m \to x_0+0} \lim_k f_k(r_m) = \lim_k f_k(x_0 + 0).$$

Since the $f_k(x)$ are quasi-continuous, the set of points E for which some $f_k(x)$ is discontinuous is denumerable. If x_0 is a point of the complementary set CE of E, then all $f_k(x)$ are continuous and so $\lim_k f_k(x_0)$ exists and we define $f(x_0) = \lim_k f_k(x_0)$. If $\{x'_m\}$ is the denumerable set of discontinuity points in E, then, because of the boundedness of the sequence $f_k(x'_m)$, there exists a subsequence $f_{k_l}(x)$ such that $\lim_l f_{k_l}(x)$ exists for all x in E and so for all x in $[a, b]$. We define $f(x)$ as this limit.

It remains to show that this convergence is uniform on $[a, b]$. For this purpose we note that because of the right equiconvergence of $f_{k_l}(x)$ at any x_0, for $e>0$, there exists d_{ex_0} such that if $0<x_1-x_0<d_{ex_0}$ and $0<x_2-x_0<d_{ex_0}$ then for all l:

$$|f_{k_l}(x_1) - f_{k_l}(x_2)| < e.$$

By taking limits as to l we also have $|f(x_1)-f(x_2)|<e$. Moreover if $l>l_{ex_1}$ we have $|f_{k_l}(x_1)-f(x_1)|<e$. Consequently for $e>0$, there exists d_{ex_0} and $l_{ex_0}=l_{ex_1}$ such that $0<x_2-x_0<d_{ex0}$ and $l>l_{ex_0}$ implies

$$|f_{k_l}(x_2) - f(x_2)| < 3e.$$

A similar statement holds to the left of each x_0. If we consider the intervals $[x_0, x_0+d]$ with $d<d_{ex_0}$ for $a \leqq x_0 < b$ and the intervals $[x_0-d, x_0]$ with $d<d'_{ex_0}$ for $a<x_0 \leqq b$, then these form sets of intervals to which the Young-Lusin theorem [**2**, pp. 441–2] applies, i.e. there exists a finite number of nonoverlapping intervals chosen from these left and right intervals reaching from a to b. Let the end points of these intervals be $a=x_0, x_1, \cdots, x_q=b$. Then if l_e is the largest of the l_{ex_i} obtained above, we will have $|f_{k_l}(x)-f(x)|<3e$ for $l>l_e$ provided x is interior to one of the intervals (x_i, x_{i+1}). Since we have convergence at each of the points x_i, we can find an l'_e which will do at these points as well as at the interior points of the intervals, i.e. we have: for $e>0$, there exists l'_e such that if $l>l'_e$ and $a \leqq x \leqq b$, then $|f_{k_l}(x)-f(x)|<3e$, which is the uniformity desired.

It now follows that $f(x)$ is in Q.

Changes to be made in this theorem when Q is replaced by the subclass Q_r of quasi-continuous functions continuous on the right are obvious.

References

1. T. H. Hildebrandt, Theory of Integration, Academic Press, New York, 1963.

2. ———, The Borel theorem, Bull. Amer. Math. Soc., 32 (1926) 423–474.

THE USE OF GREEN'S FUNCTIONS TO SOLVE SOME GENERAL INITIAL-VALUE PROBLEMS

C. E. FALBO, Sonoma State College

Ince [1] constructs a "Green's function" solution to an nth order completely homogeneous linear differential system. His solution has continuous derivatives of order $n-2$ over its entire domain, with the $(n-1)$st derivative discontinuous at some point in the domain.

In this paper, I use functions which solve initial value problems whose differential equations are (in turn) nth order linear, nth order nonlinear and homogeneous in the dependent variable. If the equation satisfies a certain "uniqueness" condition, the solution is the *only* one. The Green's functions used here are not restricted to continuity of the first $n-2$ derivatives.

One of the problems solved requires the solution to be a function with k cusp points (k is a positive integer greater than one). A nonsymmetric function is obtained as the solution to a problem in which the linear nth order differential equation is "complete," i.e., nonhomogeneous.

1. Introduction.

1.1. *Notational Definitions.* Let $a<b$, then: the interval $[a, b]$ is the set $\{x \mid a \leqq x \leqq b\}$, the segment (a, b) is the set $\{x \mid a<x<b\}$, $[a, b)$ is the set $\{x \mid a \leqq x<b\}$, and $(a, b]$ is the set $\{x \mid a<x \leqq b\}$. Let n denote a positive integer, and suppose that the nth derivative of the function f exists on the number set M. If t is in M, then $f^{(n)}(t)$ denotes the nth derivative of f at t. $f^{(n)}$ denotes the function $\{(t, f^{(n)}(t)) \mid t \text{ is in } M\}$. "$f$ is in C^n on M" means that $f^{(n)}$ is continuous on M.

$$f^{(n)}(s^+) = \lim_{\substack{t\to s\\ t>s}} f^{(n)}(t),$$

with a similar definition of $f^{(n)}(s^-)$ from the left. Let

$$W_i(f, g)_t = f(t)g^{(i)}(t) - g(t)f^{(i)}(t).$$

$\bar{f}(t)$ is the point with coordinates $(f(t), f^{(1)}(t), \cdots, f^{(n)}(t))$ of E_{n+1} (Euclidean $n+1$ space). $\bar{f}(t, s)$ is the point of E_{n+1} with coordinates $(f_1(t, s), f_1^{(1)}(t, s), \cdots, f_1^{(n)}(t, s))$ where $f_1^{(i)}(t, s)$ is the ith partial derivative of f with respect to the first variable at the point (t, s) in the domain of f.

1.2. *Green's function.* In this paper the following initial-value problem is generalized:

Let s denote a number in the segment $(0, 1)$ and suppose that f is a function such that: $f^{(2)}(t)=0$ on $[0, s)+(s, 1]$, f is continuous on $[0, 1]$ and in C^2 on $[0, s)+(s, 1]$, $f(0)=f(1)=0$, and $f^{(1)}(s^+)-f^{(1)}(s^-)=-1$. Then the only solution is the function G defined by the equation:

$$G(s, t) = \begin{cases} t(1 - s) & \text{if } t \text{ is in } [0, s] \\ s(1 - t) & \text{if } t \text{ is in } [s, 1]. \end{cases} \tag{1}$$

Such a function is usually called a Green's function. In general, the following definition is proposed. "G is a Green's function of class C^n on the domain $A = \{(x, y) \mid a \leqq x \leqq b; a \leqq y \leqq b\}$" means that G is the function $\{(x, y, G(x, y)) \mid (x, y)$ is in $A\}$ and:

1. G is continuous and symmetric on its domain, i.e. $G(x, y) = G(y, x)$.
2. For each number s in the segment (a, b), $G(t, s)$ is of class C^n for each number t in $[a, s) + (s, b]$.
3. For each positive integer $i \leqq n-1$ there is a number J_{is} such that $G_1^{(i)}(s^+, s) - G_1^{(i)}(s^-, s) = J_{is}$ and not all of the numbers J_{is} are zero.

1.3. *Types of equations.* The theorems here are concerned with linear equations and non-linear y-homogeneous equations. Definitions: "f is linear in $\bar{y}$" means that for u in C^n and v in C^n, and for a and b numbers, $f(a\bar{u}+b\bar{v}) = af(\bar{u}) + bf(\bar{v})$ "f is homogeneous in $\bar{y}$" means that there exists a number m such that $f(c\bar{y}) = c^m f(\bar{y})$. The differential equation $f[\bar{y}(t)] = g(t)$ for t on M is "complete" if $g(t) \not\equiv 0$ on M, "reduced" if $g(t) \equiv 0$ on M, "linear" if f is linear on M, "nonlinear" if f is nonlinear, and "y-homogeneous" if $g(t) \equiv 0$ on M and f is homogeneous in $\bar{y}$.

1.4. *Lemma on y-homogeneous differential equations.*

LEMMA 1. (a) *Every reduced linear differential equation is y-homogeneous.* (b) *There exists a nonlinear differential equation which is y-homogeneous.* (c) *There exists a reduced nonlinear differential equation which is not y-homogeneous.*

Part (a) follows from the definitions. For part (b) consider the function $F[\bar{y}(t)] = [y^{(3)}(t)]^2 + y(t)y^{(1)}(t)$. Here F is nonlinear and homogeneous in $\bar{y}$, and the equation $F[\bar{y}(t)] = 0$ a nonlinear y-homogeneous differential equation. For part (c), the equation:

$$y^{(1)}(t) + y^{(2)}(t)y^{(3)}(t) = 0$$

is an example of a nonlinear reduced equation which is not y-homogeneous.

1.5. *Definition of uniqueness.*

DEFINITION. *Let $s_1, s_2, \cdots, s_k$ denote a set of k numbers in the segment (a, b). The statement that "the n-th order differential equation: $F[\bar{u}(t)] = g(t)$ satisfies the uniqueness condition with respect to $s_1, \cdots, s_k$ on $[a, b]$," means that there exists one and only one function u of class C^n on $[a, b]$ such that $u(a) = u(b) = u^{(i)}(b) = 0$ for $i = 1, 2, \cdots, n-1$ satisfying the equation $F[\bar{u}(t)] = 0$ on $[a, b]$ except, possibly at $t = s_j$, $j = 1, 2, \cdots, k$.*

2. Initial-value problems involving reduced equations.

2.1. THEOREM 1. *Suppose that the n-th order linear differential equation*

$F[\bar{u}(t)]=0$ satisfies the uniqueness condition on $[a, b]$. Let α denote a class C^n solution to the system:

$$F[\bar{y}(t)] = 0$$
$$y(a) = a_0,$$

and let β denote a class C^n solution to the system:

$$F[\bar{y}(t)] = 0$$
$$y(b) = b_0.$$

Suppose further that $\beta \neq \alpha$ on $[a, b]$. (See section 1.2, considering there $\alpha(t)=t$ and $\beta(t)=(1-t)$.)

For each number s in the segment (a, b) the Green's function G defined by

$$G(t, s) = \begin{cases} \alpha(t)\beta(s) & \text{if } t \text{ is in } [a, s] \\ \beta(t)\alpha(s) & \text{if } t \text{ is in } [s, b] \end{cases} \tag{2}$$

is the only solution to the initial value problem:

$$\left.\begin{array}{ll} \text{(i)} & \textit{The function } x=\{(t, x(t, s)\} \textit{ is continuous for } t \textit{ in } [a, b] \textit{ and of class } C^n \\ & \textit{on } [a, s)+(s, b], \\ \text{(ii)} & F[\bar{x}(t, s)]=0 \textit{ on } [a, s)+(s, b],\ x(a, s)=a_0\beta(s),\ x(b, s)=b_0\alpha(s), \\ \text{(iii)} & x^{(i)}(s^+, s) \textit{ exists and } x^{(i)}(s^-, s) \textit{ exists, and } x^{(i)}(s^+, s)-x^{(i)}(s^-, s) \\ & =W_i(\alpha, \beta)_s\ i=1, 2, \cdots, n-1, \textit{ and} \\ \text{(iv)} & x^{(i)}(b, s)=\alpha(s)\cdot\beta^{(i)}(b),\ i=1, 2, \cdots, n-1. \end{array}\right\} \tag{3}$$

Proof. The function G defined in (2) satisfies (3). That G satisfies (i), (ii) and (iv) of (3) is clear. A proof that G satisfies (iii) is as follows: $G^{(i)}(s^+, s)$ exists [and is equal to $\alpha(s)\cdot\beta^{(i)}(s)$]. For if $\alpha(s)=0$, then $\alpha(s)\cdot\beta^{(i)}(t)=\alpha(s)\beta^{(i)}(s)=0$, hence

$$\lim_{\substack{t\to s \\ t>s}} \alpha(s)\beta^{(i)}(t) = \alpha(s)\beta^{(i)}(s),$$

therefore $G^{(i)}(s^+, s)$ exists and is equal to $\alpha(s)\beta^{(i)}(s)$. Suppose $\alpha(s)\neq 0$. Since β is in C^n on $[a, b]$, then for $i=1, 2, \cdots, n$, $\beta^{(i)}$ is continuous at s. By definition of continuity, if $c>0$, there exists a positive number d' such that if $0<d<$the minimum of d' and $b-s$, then for each number t such that $0<t-s<d$,

$$|\beta^{(i)}(s) - \beta^{(i)}(t)| < \frac{c}{|\alpha(s)| + 1},$$

multiplying this inequality by $|\alpha(s)|>0$, we have

$$|\beta^{(i)}(s) - \beta^{(i)}(t)| \cdot |\alpha(s)| < \frac{c}{|\alpha(s)| + 1} \cdot |\alpha(s)| < c.$$

Therefore

$$|\beta^{(i)}(s)\alpha(s) - \beta^{(i)}(t)\alpha(s)| < c. \quad \text{or}$$

$$|\beta^{(i)}(s)\alpha(s) - G^{(i)}(t, s)| < c \quad \text{for } s < t \leqq b.$$

Thus $\lim_{t\to s, t>s} G^{(i)}(t, s)$ exists, i.e. $G^{(i)}(s^+, s)$ exists and is $\beta^{(i)}(s)\alpha(s)$. Similarly, $G^{(i)}(s^-, s)$ exists and is $\alpha^{(i)}(s)\beta(s)$. We now see that G satisfies (iii).

Granting the conditions in (3), it will be shown that (2) follows. Suppose that z is a function satisfying (3), and G is the function defined in (2). For each number t in $[a, b]$, let $e(t, s) = G(t, s) - z(t, s)$. Then:

$$\left.\begin{array}{ll}
\text{(i)}' & e \text{ is continuous on } [a, b], \text{ and of class } C^n \text{ in } t \text{ on } [a, s)+(s, b], \\
\text{(ii)}' & F[\bar{e}(t, s)] = 0 \text{ on } [a, s)+(s, b],\ e(a, s) = e(b, s) = 0, \\
\text{(iii)}' & e^{(i)}(s^+, s) \text{ exists and } e^{(i)}(s^-, s) \text{ exists, and } e^{(i)}(s^+, s) - e^{(i)}(s^-, s) = 0, \\
& i \leqq n-1, \text{ and} \\
\text{(iv)}' & e^{(i)}(b, s) = 0,\ i \leqq n-1.
\end{array}\right\} \quad (3)'$$

The system (3)′ follows immediately from the definition of e and from the system (3). We shall prove that e is of class C^n on the entire interval $[a, b]$, then using the uniqueness condition, we have that $e \equiv 0$ on $[a, b]$, i.e. the function G of (2) is the only solution to (3).

Suppose that there exists a positive number c such that if h is any positive number, then there exists a number t_1 in the segment $(s-h, s)$ such that

$$\left|\frac{e(s, s) - e(t_1, s)}{s - t_1} - e^{(1)}(s^-, s)\right| \geqq c. \tag{4}$$

Since e is of class C^n in $(s-h, s)$ and since (t_1, s) is a subsegment of $(s-h, s)$, then e is of class C^n in (t_1, s), and e is continuous in the interval $[t_1, s]$. Therefore, by the law of the mean, there is a number t^* in (t_1, s) such that

$$\frac{e(s, s) - e(t_1, s)}{s - t_1} = e^{(1)}(t^*, s). \tag{5}$$

Substituting from equation (5) into inequality (4) we get

$$|e^{(1)}(t^*, s) - e^{(1)}(s^-, s)| \geqq c. \tag{6}$$

This means that there exists a positive number c such that if h is any positive number, then there exists a number t^* in the segment $(s-h, s)$ such that inequality (6) holds, which is contrary to the definition:

$$\operatorname*{Lim}_{\substack{t\to s \\ t<s}} e^{(1)}(t, s) = e^{(1)}(s^-, s).$$

Therefore, the supposition is false and we may state its denial. If $c > 0$, then there exists a positive number d', such that if $0 < d_1 <$ the minimum of d' and $(s-a)/2$,

and t is a number such that $0<s-t<d_1$, then

$$\left|\frac{e(s, s) - e(t, s)}{s - t} - e^{(1)}(s^-, s)\right| < c. \tag{7}$$

Similarly, for $t>s$, there exists a positive number d'', such that if $0<d_2<$the minimum of d'' and $(b-s)/2$ and t is a number such that $0<t-s<d_2$, then

$$\left|\frac{e(s, s) - e(t, s)}{s - t} - e^{(1)}(s^+, s)\right| < c \tag{8}$$

Hence $e^{(1)}(s, s)$ exists and is equal to $e^{(1)}(s^-, s)$ [also to $e^{(1)}(s^+, s)$]. This conclusion with condition (i)′ proves that e is of class C^1 on the entire interval $[a, b]$.

Suppose that k is a positive integer, and $k\leqq n-1$, and e is of class C^k on $[a, b]$, then $e^{(k+1)}(s, s)$ exists. The proof of this follows the same arguments as above, and is omitted here. The conclusion is that $e^{(k+1)}(s, s)$ exists and is equal to $e^{(k+1)}(s^-, s)$ and to $e^{(k+1)}(s^+, s)$. Thus e is of class C^{k+1} on the entire interval $[a, b]$, and this is true for $k=1, 2, \cdots, n-1$: therefore e is in C^n on $[a, b]$.

We now know that e is a function of class C^n on $[a, b]$ with the properties:

$$e(a, s) = e(b, s) = e^{(i)}(b, s) = 0, \quad \text{for } i = 1, 2, \cdots, n \tag{9}$$

$$F[\bar{e}(t, s)] = 0 \quad \text{on } [a, s) + (s, b]. \tag{10}$$

Since the differential equation of this theorem satisfies the uniqueness condition on $[a, b]$, then there is only one function e such that (9) and (10) are true. We note that these equations are also true for $p(t, s)=0$ on $[a, b]$. Because of the uniqueness condition we have that $e(t, s)=p(t, s)=0$ on $[a, s)+(s, b]$. Since e is continuous at s, we have $e(s, s)=0$. This proves that (2) is the only solution to (3).

2.2. Corollary 1 to Theorem 1. *If the differential equation*:

$$F[\bar{y}(t, s)] = 0 \tag{11}$$

is nonlinear and y-homogeneous, then the function G defined in (2) *solves the initial-value problem* (3) *with equation* (11) *in place of* (ii).

Proof. That a nonlinear y-homogeneous differential equation exists is proved in Lemma 1; hence this corollary is not vacuous. There exists a number p such that, if c is any number, then $F[c\bar{y}(t, s)]=c^pF[\bar{y}(t, s)]$, because F is y-homogeneous. Let α and β satisfy (11) on $[a, b]$, then $c_1\alpha$ and $c_2\beta$, for any numbers c_1 and c_2, satisfy (11) on $[a, b]$. Hence, for any number s in the segment (a, b),$\beta(s)\alpha$ and $\alpha(s)\beta$ satisfy (11).

If the function G is as defined in (2) then, from the proof of Theorem 1, G is continuous on $[a, b]$ and of class C^n on $[a, s)+(s, b]$. Each of $G^{(i)}(s^+, s)$ and $G^{(i)}(s^-, s)$ exists, and $G^{(i)}(s^+, s)-G^{(i)}(s^-, s)=W_i(\alpha, \beta)_s$. If the initial conditions on α and β are $\alpha(a)=a_0$, and $\beta(b)=b_0$, then it follows that (2) solves the initial-value problem (3) for the nonlinear aifferential equation (11). In Theorem 1,

the uniqueness condition was used in proving that there is no other solution to (3), and is therefore not necessary here.

2.3. COROLLARY 2 TO THEOREM 1. *Assume the conditions preceding* (2) *in the theorem. If, in addition, the term containing y in the linear differential equation has zero for its coefficient, then the function G defined by*

$$G(t, s) = \begin{cases} \alpha(t) + \beta(s) & \text{if } t \text{ is in } [a, s] \\ \beta(t) + \alpha(s) & \text{if } t \text{ is in } [s, b] \end{cases} \tag{12}$$

is the only solution to the initial-value problem:

$$\left.\begin{array}{ll} \text{(i)} & \textit{The function } x = \{(t, x(t, s)\} \textit{ is continuous for } t \textit{ in } [a, b] \textit{ and of class} \\ & C^n \textit{ on } [a, s) + (s, b], \\ \text{(ii)} & F[\bar{x}(t, s)] = 0 \textit{ on } [a, s) + (s, b],\ x(a, s) = a_0 + \beta(s),\ x(b, s) = b_0 + \alpha(s), \\ \text{(iii)} & \textit{For each positive integer } i \leqq n-1, \\ & x^{(i)}(s^+, s) \textit{ and } x^{(i)}(s^-, s) \textit{ exist}, \\ & x^{(i)}(s^+, s) - x^{(i)}(s^-, s) = \beta^{(i)}(s) - \alpha^{(i)}(s), \textit{ and} \\ \text{(iv)} & x^{(i)}(b, s) = \beta^{(i)}(b),\ i = 1, 2, \cdots, n-1. \end{array}\right\} \tag{13}$$

Proof. Here F is linear and the proof follows similar arguments as in the proof of the theorem. The initial value problem (13) differs from (3) in that here the y term is missing, and the discontinuity described in (iii) of (13) is different from the one in (iii) of (3). The coefficient of y is zero so that $F[\alpha(t) + \beta(s)] = F[\alpha(t)] = 0$.

2.4 The case of more than one "cusp." The theorem of this section deals with a function whose graph has k cusp points in the given interval. (k is a positive integer >1). The order of the differential equation is $n \geqq 2$.

THEOREM 2. *Suppose that each of α and β is a solution to the linear n-th order differential equation*

$$F[\bar{y}(t)] = 0 \quad on \quad [0, 1], \tag{14}$$

$\alpha(0) = \beta(1) = 0$, *and the equation* (14) *satisfies the uniqueness condition on* $[0, 1]$. *Let* $r = (s_1, s_2, \cdots, s_k)$ *be a set of k numbers in the segment* $(0, 1)$, *such that* $0 < s_1 < s_2 < \cdots < s_k < 1$. *The function G defined by*:

$$G(r, t) = \begin{cases} \alpha(t) \sum_{m=1}^{m=k} \beta(s_m) & \text{if } t \text{ is in } [0, s_1]. \\ \beta(t) \sum_{m=1}^{m=j} \alpha(s_m) + \alpha(t) \sum_{m=j+1}^{m=k} \beta(s_m) & \text{if } t \text{ is in } [s_j, s_{j+1}], \\ & j = 1, 2, \cdots, k-1, \\ \beta(t) \sum_{m=1}^{m=k} \alpha(s_m) & \text{if } t \text{ is in } [s_k, 1], \end{cases} \tag{15}$$

is the only solution to the initial value problem:

$$\left.\begin{array}{ll}\text{(i)} & \textit{The function } x=\{(t,\,x(r,t)) \textit{ for } t \textit{ in the interval } [0,1]\}, \textit{ is continuous} \\ & \textit{in } t \textit{ on } [0,1], \textit{ and } x \textit{ is of class } C^n \textit{ in } t \textit{ on } M=[0,s_1)+(s_1,s_2)+\cdots \\ & +(s_{k-1},s_k)+(s_k,1]. \\ \text{(ii)} & F[\bar{x}(r,t)]=0 \textit{ on } M. \textit{ (Here } \bar{x}(r,t) \textit{ is the point } (x(r,t),\cdots,x_t^{(i)}(r,t), \\ & \cdots, x_t^{(n)}(r,t) \textit{ of } E_{n+1}.) \ x(r,0)=x(r,1)=0. \\ \text{(iii)} & \textit{For each positive integer } i\leqq n,\ x^{(i)}(r,s_j^+) \textit{ and } x^{(i)}(r,s_j^-) \textit{ exist, and their} \\ & \textit{difference is } W_i(\alpha,\beta)_{s_j}, \textit{ for } j=1,2,\cdots,k \textit{ and} \\ \text{(iv)} & x^{(i)}(r,1)=\beta^{(i)}(1)\sum_{j=1}^{k}\alpha(s_j). \end{array}\right\} \quad (16)$$

Proof. First it is shown that (15) is a solution to (16). Assume G as defined by (15); then (i)′ G is continuous for t in $[0, 1]$ and G is of class C^n on M. This is true since, in each segment (s_j, s_{j+1}) and in $[0, s_1)$ and $(s_k, 1]$ of M, G is a linear combination of functions of class C^n. Therefore G is of class C^n on M. From this, G is continuous on M. To show that G is continuous on $[0, 1]$, all that remains is to show that G is continuous at s_j for each integer j. We settle the case for $2\leqq j\leqq k-1$. Suppose:

$$\left|\sum_{m=1}^{j}\alpha(s_m)\right| > 0 \quad\text{and}\quad \left|\sum_{m=j+1}^{k}\beta(s_m)\right| > 0.$$

Since each of α and β is continuous at s_j, then by the definition of continuity, if $\epsilon > 0$, there is a number $\delta' > 0$ such that if $0 < \delta_1 <$ minimum of δ', and $(s_j-s_{j-1})/2$ and if t is a number such that $0<s_j-t<\delta_1$, then

$$|\beta(t)-\beta(s_j)| < \frac{\epsilon/2}{\left|\sum_{m=1}^{j}\alpha(s_m)\right|+1} \tag{17}$$

and

$$|\alpha(t)-\alpha(s_j)| < \frac{\epsilon/2}{\left|\sum_{m=j+1}^{k}\beta(s_m)\right|+1} \tag{18}$$

then

$$|\beta(t)-\beta(s_j)|\cdot\left|\sum_{m=1}^{j}\alpha(s_m)\right| + |\alpha(t)-\alpha(s_j)|\cdot\left|\sum_{m=j+1}^{k}\beta(s_m)\right| < \epsilon. \tag{19}$$

Therefore

$$\left|[\beta(t)-\beta(s_j)]\sum_{m=1}^{j}\alpha(s_m) + [\alpha(t)-\alpha(s_j)]\sum_{m=j+1}^{k}\beta(s_m)\right| < \epsilon. \tag{20}$$

$$\left|\beta(t)\sum_{m=1}^{j}\alpha(s_m)+\alpha(t)\sum_{m=j+1}^{k}\beta(s_m)-\left[\beta(s_j)\sum_{m=1}^{j}\alpha(s_m)+\alpha(s_j)\sum_{m=j+1}^{k}\beta(s_m)\right]\right| < \epsilon. \tag{21}$$

Hence,

$$\left| G(r, t) - G(r, s_j) \right| < \epsilon \quad \text{if } 0 < s_j - t < \delta_1. \tag{22}$$

Similarly, there is a positive number δ'' such that if $0<\delta_2$ minimum of δ'' and $(s_{j+1}-s_j)/2$ and $0<t-s_j<\delta_2$, then

$$\left| G(r, t) - G(r, s_j) \right| < \epsilon \quad \text{if } 0 < t - s_j < \delta_2. \tag{23}$$

Suppose that either

$$\sum_{m=1}^{j} \alpha(s_m) = 0 \quad \text{or} \quad \sum_{m=j+1}^{k} \beta(s_m) = 0.$$

Consider the case in which only the first sum is zero. Then both of (17) and (18) still hold, and

$$\left| \alpha(t) - \alpha(s_j) \right| \cdot \left| \sum_{m=j+1}^{k} \beta(s_m) \right| < \epsilon.$$

Therefore (19) still holds, so does (20) and therefore (21), hence the same result is obtained. Similarly for the case in which only the second sum is zero.

If both sums are zero, then (19) still holds, etc. From (22) and (23) we see that there exists a number δ, $0<\delta<$minimum of δ_1 and δ_2 such that if $|t-s_j|<\delta$, then

$$\left| G(r, t) - G(r, s_j) \right| < \epsilon.$$

Therefore G is continuous at s_j for $2 \leqq j \leqq k-1$. By a similar argument G is continuous at s_1 and at s_k. This concludes the proof that G is continuous on $[0, 1]$.

(ii)′ $F[\overline{G}(r, t)]=0$ on M.

Proof. Since F is linear, and since each of the sums $\sum_{m=1}^{j} \alpha(s_m)$ and $\sum_{m=j}^{k} \beta(s_m)$ is a number, then on $[0, s_1)$:

$$F[\overline{G}(r, t)] = F\left(\sum_{m=1}^{k} \beta(s_m)\bar{\alpha}(t) \right) = \sum_{m=1}^{k} \beta(s_m) F[\bar{\alpha}(t)] = 0.$$

On (s_j, s_{j+1}):

$$\begin{aligned} F[\overline{G}(r, t)] &= F\left(\sum_{m=1}^{j} \alpha(s_m)\bar{\beta}(t) + \sum_{m=j+1}^{k} \beta(s_m)\bar{\alpha}(t) \right) \\ &= \sum_{m=1}^{j} \alpha(s_m) F[\bar{\beta}(t)] + \sum_{m=j+1}^{k} \beta(s_m) F[\bar{\alpha}(t)] = 0. \end{aligned}$$

Similarly, on $(s_k, 1]$: $F[\overline{G}(r, t)]=0$. Furthermore, $G(r, 0)=0$, and $G(r, 1)=0$.

By arguments similar to the ones used in the proof of Theorem 1, $G^{(i)}(r, s_j^+)$ exists, and is

$$\beta^{(i)}(s_j)\sum_{m=1}^{j}\alpha(s_m) + \alpha(s_j)\sum_{m=j+1}^{k}\beta(s_m),$$

for t in (s_j, s_{j+1}), $1 \leqq j \leqq k-1$ and $G^{(i)}(r, s_j^-)$ exists and is

$$\beta^{(i)}(s_j)\sum_{m=1}^{j-1}\alpha(s_m) + \alpha^{(i)}(s_j)\sum_{m=j}^{k}\beta(s_m),$$

for t in (s_{j-1}, s_j) and $2 \leqq j \leqq k$.

$$G^{(i)}(r, s_1^-) = \alpha^{(i)}(s_1)\sum_{m=1}^{k}\beta(s_m), \quad \text{and} \quad G^{(i)}(r, s_k^+) = \beta^{(i)}(s_k)\sum_{m=1}^{k}\alpha(s_m).$$

Therefore $G^{(i)}(r, s_j^+) - G^{(i)}(r, s_j^-) = W_i(\alpha, \beta)_{s_j}$. For example, if $1 < j < k$, then

$$\begin{aligned} G^{(i)}(r, s_j^+) - G^{(i)}(r, s_j^-) &= \beta^{(i)}(s_j)\sum_{m=1}^{j}\alpha(s_m) - \beta^{(i)}(s_j)\sum_{m=1}^{j-1}\alpha(s_m) \\ &\quad + \alpha^{(i)}(s_j)\sum_{m=j+1}^{k}\beta(s_m) - \alpha^{(i)}(s_j)\sum_{m=j}^{k}\beta(s_m) \\ &= \beta^{(i)}(s_j)\alpha(s_j) - \alpha^{(i)}(s_j)\beta(s_j) \\ &= W_i(\alpha, \beta)s_j. \end{aligned}$$

To complete the proof that (15) satisfies (16), we have the result

(iv) $$G^{(i)}(r, 1) = \beta^{(i)}(1)\sum_{m=1}^{k}\alpha(s_m).$$

Now to prove that (15) is the only solution to (16). Let G denote the function defined in (15). If z is a function satisfying (16), let e be the function defined by $e(r, t) = G(r, t) - z(r, t)$; r is as before and t is in $[0, 1]$. e satisfies the conditions:

(a)′ e is continuous on $[0, 1]$, and of class C^n on M.

(b)′ $F[\bar{e}(r, t)] = 0$, on M.

(c)′ $e(r, 0) = e(r, 1) = 0$.

(d)′ For $i \leqq n-1$, $e^{(i)}(r, s_j) - e^{(i)}(r, s_j) = 0$, and $e^{(i)}(r, 1) = 0$.

By an argument similar to that in the proof of Theorem 1, e is of class C^n on $[0, 1]$, and by the uniqueness condition $e(r, t)$ is the only function satisfying (a)′ through (d)′, but the function $p(r, t) \equiv 0$ also satisfies these conditions; hence $e(r, t) \equiv 0$ on $[0, 1]$, and (15) is the only solution to (16).

3. The Complete Equation.

3.1. LEMMA 2. *There exists a positive integer n, a linear function L from a do-*

main in E_{n+1}, *a function* f *on* $[0, 1]$, *and two functions* α *on* $[0, 1]$, *and* β *on the set* $A=\{(x, y) \mid 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1\}$ *satisfying*

$$\begin{aligned} \beta(s, s) L[\bar{\alpha}(t)] &= f(t) \\ \alpha(s) L[\bar{\beta}(t, s)] &= f(t) \end{aligned} \tag{24}$$

and $\alpha(0)=\beta(1, s)=0$ *for each number* s *in the segment* $(0, 1)$.

Proof. Let $n=2$, $L[\bar{y}(t)]=y^{(2)}(t)$, $f(t)=6t$, then if $\alpha(t)=t^3+t$ and

$$\beta(t, s)=\frac{t^3-1}{s^3+s}+\frac{(s+1)(t-1)}{(s^3+s)(s-1)}, \quad \text{then} \quad \alpha(s) \beta_t^{(2)}(t, s)=6t,$$

$\beta(s, s) \alpha^{(2)}(t)=6t$, and $\alpha(0)=\beta(1, s)=0$, thus (24) is satisfied.

3.2. THEOREM 3. *Suppose that* L *is a linear function from a domain in* E_{n+1}, *each of* f *and* α *is a function on* $[0, 1]$, *and* β *is a function on the set* $A=\{(x, y) \mid 0 \leqq x \leqq 1,\ 0 \leqq y \leqq 1\}$ *such that*

$$\begin{aligned} \beta(s, s) L[\bar{\alpha}(t)] &= f(t) \\ \alpha(s) L[\bar{\beta}(t, s)] &= f(t) \end{aligned} \tag{25}$$

and $\alpha(0)=\beta(1, s)=0$ *for each number* s *in the segment* $(0, 1)$. *Suppose, further, that each equation in* (25) *satisfies the uniqueness condition on* $[0, 1]$. *The function* G *defined by*

$$G(t, s)=\begin{cases} \alpha(t) \beta(s, s) & \text{for } t \text{ in } [0, s] \\ \beta(t, s) \alpha(s) & \text{for } t \text{ in } [s, 1] \end{cases} \tag{26}$$

is the only solution to the initial value problem:

(i) x *is continuous in* t *on* $[0, 1]$, *and of class* C^n *in* t *on* $[0, s)+(s, 1]$.

(ii) $L[\bar{x}(t, s)]=f(t)$ *on* $[0, s)+(s, 1]$, $x(0, s)=x(1, s)=0$.

(iii) *For each positive integer* $i \leqq n-1$, $x^{(i)}(s^+, s)$ *and* $x^{(i)}(s^-, s)$ *exist, and their difference is* $w_i(\alpha, \beta)_s$. (27)

(iv) $x^{(i)}(1, s)=\alpha(s) \beta^{(i)}(1, s)$.

Proof. It is clear that the function defined in (26) is a solution to (27). To show that (27) implies (26), suppose that z is a function such that z satisfies (27), and let $e(s, t)=G(s, t)-z(s, t)$, where G is the function defined in (26), then

(a)′ e is continuous on $[0, 1]$, and of class C^n in t on $[0, s)+(s, 1]$.

(b)′ $L[\bar{e}(t, s)]=0$ on $[0, s)+(s, 1]$.

(c)′ $e(0, s)=e(1, s)=0$.

(d)′ $e^{(i)}(s^+, s)-e^{(i)}(s^-, s)=0$.

Also $e^{(i)}(1, s)=0$. By an argument similar to the one used in the proof of Theorem 1, e is of class C^n on $[0, 1]$. By the uniqueness condition e is the only function satisfying (a)′ through (d)′. Zero also satisfies these conditions. Therefore $e(t, s) \equiv 0$ on $[0, 1]$ and G as defined in (26) is the only solution.

4. An application. If each of P_0 and P_1 is a continuous function on $[0, 1]$, if α and β are solutions to

$$x^{(2)}(t) + P_1(t)x^{(1)}(t) + P_0(t)x(t) = 0, \text{ on } [0, 1] \tag{28}$$

with $\alpha(0) = \beta(1) = 0$, if g is integrable, and if

$$K(s, t) = \begin{cases} \alpha(t)\beta(s) & \text{for } t \text{ in } [0, s] \\ \beta(t)\alpha(s) & \text{for } t \text{ in } [s, 1] \end{cases}$$

then the integral

$$y(t) = \int_0^1 K(s, t)g(s)\,ds \tag{29}$$

is the only solution to the problem

$$\begin{gathered} y^{(2)}(t) + P_1(t)y^{(1)}(t) + P_0(t)y(t) = g(t)W_1(\alpha, \beta)_t \\ y(0) = y(1) = 0. \end{gathered} \tag{30}$$

As a specific example, consider: $\alpha(t) = a \sin(t)$, and $\beta(t) = b \cos(t)$ which are solutions to

$$x + x'' = 0 \quad \text{with} \quad \alpha(0) = \beta(\pi/2) = 0.$$

Then the integral:

$$y(t) = \int_0^{\pi/2} G(s, t)g(s)\,ds, \quad g \text{ in } C^1 \quad \text{and}$$

$$G(s, t) = \begin{cases} a \sin(t)\, b \cos(s) & \text{for } t \text{ in } [0, s] \\ a \cos(t)\, a \sin(s) & \text{for } t \text{ in } [s, \pi/2] \end{cases}$$

is the only solution to:

$$\begin{gathered} y + y'' = -abg \\ y(0) = y(\pi/2) = 0. \end{gathered}$$

The use of Green's function as a kernel of an integral can be applied to higher order equations, as in the following example. If each of P_0, P_1, P_2 is of class C^0 on $[a, b]$, let

$$F[x(t)] = x^{(3)}(t) + P_2(t)x^{(2)}(t) + P_1(t)x^{(1)}(t) + P_0(t)x(t).$$

If the equation $F[x(t)] = 0$ satisfies the uniqueness condition with respect to s, s in the segment (a, b), and if g is in C^2 and α and β are two (distinct) C^3 solutions to $F[x(t)] = 0$ with $\alpha(a) = \beta(b) = \beta^{(i)}(b) = 0$, $i = 1, 2$. Then the integral:

$$\left.\begin{gathered} y(t) = \int_a^b G(s, t)g(s)\,ds \\ G \text{ is as defined in equation (2)} \end{gathered}\right\} \tag{31}$$

is the only solution to:

$$F[y(t)] = g(t)[P_2(t)W_1(\alpha,\beta)_t + 2W_2(\alpha,\beta)_t] + g'(t)W_1(\alpha,\beta)_t \qquad (32)$$
$$y(a) = y(b) = y'(b) = 0.$$

To prove that (31) is a solution to (32), write $y(t)$ as:

$$y(t) = \int_a^t G(s,t)g(s)\,ds + \int_t^b G(s,t)g(s)\,ds = \int_a^t \beta(t)\alpha(s)g(s)\,ds$$
$$+ \int_t^b \alpha(t)\beta(s)g(s)\,ds.$$

Using the integral $y(t) = \int_a^b G(s, t)g(s)ds$, g in C^{n-1} as a solution to the nth order linear differential equation analogous to (32), the Green's function of Theorem 1 yields a result, but it is quite cumbersome. For example, the first term alone (i.e. $y^{(n)}(t)$) produces:

$$\beta^{(n)}(t)\int_a^t \alpha(s)g(s)\,ds + \alpha^{(n)}(t)\int_t^b \beta(s)g(s)\,ds$$
$$+ \sum_{i=1}^{n}\binom{n}{i}\sum_{j=0}^{i-1}\binom{i-1}{j} g^{(i-1-j)}(t)W_{n-i-j}(\alpha^{(j)},\beta^{(j)})t.$$

When $W_n(f, g)$ is as defined before for n positive, is zero for $n=0$, and is defined by $W_{-k}(f, g) = fD^{-k}f - gD^{-k}f$, for $-k$ negative, where $D^{-k} = k$th antiderivative. Of course, in the rest of the equation the sum of the terms containing the integrals is zero, while the other terms will be somewhat simplified.

The author is grateful to Professor H. J. Ettlinger for several helpful suggestions.

Reference

1. E. L. Ince. Ordinary Differential Equations, Dover, New York, 1956, pp. 254–261.

THE MATHEMATICAL ASSOCIATION OF AMERICA

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN
MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 5

CONTENTS

MAY 1966

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, N. Y. 14214.

NOTICE TO AUTHORS

More than enough Main Articles, Mathematical Notes, and Classroom Notes have been accepted (or are under consideration) to fill these Sections of the MONTHLY through the December, 1966 issue. Please do not submit any papers for these Sections until after July 1. The June–July issue will contain addresses to which papers may then be submitted.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

PRINTED IN THE UNITED STATES OF AMERICA

REFORM IN TEACHING MATHEMATICS

ROLF NEVANLINNA, Academy of Finland

The question of teaching mathematics at high-school and university level has been debated ever since Felix Klein, at the end of the 19th century, presented his famous Erlanger program. Klein wanted to enliven teaching by giving prominence, in all fields of exact science—analysis, geometry and physics—to the concept of a group and to the principles of transformations. The aim of mathematical theory is that of uniting and studying the invariant elements which remain unchanged during group transformations. These invariant laws constitute the heart of the theory, what is often referred to as its "mathematical structure."

Naturally enough, Klein, who contributed greatly to the development of group theory, was inclined to overemphasize to some extent the importance of the group concept. In fact, during the subsequent progress of science, this concept has become a special case of the more general concept of mapping, which, influenced by the epoch-making ideas of Gauss and Riemann, is assuming an increasingly dominant position in mathematical science today.

With this reservation, it can be agreed that Klein in his program clearly laid down a main line for present and future mathematical science. His program has contributed essentially to the development of mathematical teaching at the university level.

These views of Klein must be given less prominence in the high school curriculum. Here, Klein's most important contribution takes another direction. He considered the time was ripe for an extension of high school mathematics so as to include the calculus, the basis of "higher mathematics." The idea has gradually gained ground. Nowadays, the syllabus in almost all countries includes a section on the concept of limit, and the elements of infinitesimal calculus.

The extent to which this involves actual progress is a controversial matter. In the experience of many university teachers, the inadequate knowledge of the calculus gained by mathematics students during their high school years seems rather to hamper them at university level, at which students are prepared for more advanced mathematics at a standard of strictness which is quite different. This sceptical attitude is not entirely unjustified. Nevertheless, a return to the former state of affairs is neither possible nor desirable. By virtue of the constantly growing significance of mathematics as a fundamental technical instrument in many different fields where the chief emphasis is laid on the application of the mathematical calculus—rather than on complete mastery of the logical structure of the mathematical technique—it is impossible to deny the need for certain, if no more than superficial, knowledge of the first principles of the infinitesimal calculus before and on university entrance.

Since the beginning of this century, mathematics has expanded enormously in various directions. Modern scientific progress is based essentially upon the

more profound understanding of the nature of mathematical knowledge attained by the so-called "axiomatic school."

David Hilbert's contribution in the early years of this century was decisive in this respect. His investigations into the foundations of geometry led to further elucidation of the logical connections upon which the elementary geometrical systems are based. But his exposition aimed higher. The enormous significance of Hilbert's critical elementary geometrical studies is chiefly the outcome of the general *fundamental attitude* he adopted towards the geometrical problems occupying his mind. This even sets its stamp upon his noteworthy sentences in the introduction to his presentation of Euclidean geometry.

Let us assume three kinds of objects, which we shall call points, lines and planes. Let us further assume that these objects are linked by certain *relations*.

Hilbert then states the three basic relations of Euclidean geometry: the relation of incidence, the relation of correspondence, and the relation of congruence.

Now follows the formulation of the axioms and, thereby, Euclidean geometry is exhibited as a whole, though in implicit form. It is the subsequent logical synthesis that leads from the basic concepts (objects, relations and rules) to new objects and relations (through definitions) and new rules or theorems (through deductions or proofs).

Hilbert's presentation is linked to the earlier discussion on geometry which, especially in the 19th century, led to the achievement of decisively new knowledge (Gauss, Riemann, Klein, Helmholtz, Poincaré, et al.). But more clearly and explicitly than anyone before him, Hilbert emphasized that a full knowledge of the nature of geometry can be obtained only if once and for all there is left out of account those "geometrical qualities" of the geometrical objects, and the relations originally associated with them, on the intuitive-empirical level of geometry. "Points," "lines," and so on, become dissociated from these conceptions, which are irrelevant from the logical point of view; they are understood only as abstract things, as elements of certain given sets.

Thus the geometrical relations are also dissociated from the intuitive geometrical views originally attached to them. There is required no great freedom of thought or capacity for abstraction to comprehend and appreciate the significance of this process of dissociation, inherent in Hilbert's conception as compared with the older, more traditional intuitive presentations of the nature of geometrical objects, once the thought has been expressed with sufficient clarity and weight, and illustrated by qualitatively different "reinterpretations," primarily but not solely by the arithmetical models represented by analytical geometry. This applies at least to the conception of the basic geometrical objects (points, lines, planes) as elements of three kinds of sets without other more special graphic qualities than those presupposed by the "intuitive" set concept: a set is a collection of given elements without qualitative characteristics other than "individual existence," so that two arbitrary elements are either identical ("same elements") or nonidentical ("different elements").

Perhaps the abstract concept of *relation* presents more difficulty. It seems that, even among specialists interested in philosophy, the word "relation" is occasionally combined with vague ideas. For a mathematician, the matter is as clear as it is simple. What is decisive is the *idea of correspondence* (or more generally the mapping idea). For example, the fact that two elementary geometrical sets, points and lines, are linked by an incidence relation means no more than that for every point there is a corresponding set of lines (which are said to be incident with the given point).

The transition from the original intuitive idea of the meaning of the incidence relation ("a line passes through a point") to the wider meaning based upon the more general concept of "correspondence" has an earlier equivalent in the logical liberation process undergone by the *concept of function* during the last century. Originally, this concept was defined by an arithmetical or analytical law which, in the case of a given number or argument, yielded as "result" another number, the value of the function (or many other numbers), if the function was multivalued. The modern idea of a function disregards the arithmetical-analytical manipulations which lead from the argument to the function. These have as substitute the general correspondence or mapping concept: an exceedingly significant insight which in all essentials goes back to Riemann. A similar process also caused the intuitive concept of "projection" to emerge as the general idea of mapping and abstract correspondence.

Hilbert's contribution represents one of the most notable stages in the history of scientific ideas. It is most remarkable that the new general insight into the nature of the formation of mathematical theory was achieved as the final result of scientists' efforts to clarify the oldest and most elementary problems of exact research: Euclidean geometry.

The free and general conception held by Hilbert on the system of elementary geometry also led him to pose the general questions on knowledge associated with *every* mathematical theory which reaches the axiomatic level, so that the theory emerges as a logical "theory of structure," as a system of basic objects, basic relations and basic sets (axioms) assumed to be valid between these objects and relations. These questions are concerned with (1) the consistency of the axiom system, (2) the independence of the system, and (3) the completeness of the system.

With the aid of analytical geometry these three questions are reduced, as regards both Euclidean and hyperbolic elementary geometry, to the question of whether *arithmetic constitutes a consistent logical system.* If this is assumed, then it follows that:

(1) The Euclidean axiom system is also consistent: it cannot lead to two consequences involving an explicit contradiction.

(2) The Euclidean axioms can be reduced so as to be mutually independent logically.

(3) Euclidean geometry is a complete system. This means that its logical structure is uniquely determined apart from isomorphy: two "realizations" or "models" of Euclidean geometry are isomorphic if their basic objects and basic relations can be mapped in one-to-one correspondence, with their relations remaining valid (two objects in one model which are linked to an elementary geometrical basic relation have, as image objects in the other model, two objects linked to the "image relation" of the relation concerned).

The question of the logical nature of geometry thus led Hilbert to a corresponding problem in arithmetic. Discussions on these questions have been in progress for half a century, and are far from being completed today.

The break-through of the "axiomatic model of thought" has not only left its imprint on research into the foundation of mathematics, but has also decisively influenced the whole expansive trend of mathematics in our time. A non-mathematician may be astonished at this statement. Does not the nature of mathematics as such imply that its method is axiomatic? What is the aim of mathematics, and what can it be, but to draw correct conclusions from clearly formulated premises (axioms) through logically valid deductions? This objection is warranted and it may be questioned whether the title "axiomatic method" properly describes what is *most essential* in this connection.

In fact, what is new in the axiomatic mode of thought is concerned less with the more stringent demands for logical exactness imposed upon mathematics today than with the freer, more abstract interpretation of the nature of mathematical theories to which the liberation discussed above has led as a result of fundamental studies in elementary geometry. I here call to mind only two main fields where the axiomatic mode of thought has brought about a thorough, large-scale renewal and fruitful expansion: "modern" algebra and topology.

During recent decades, the French Bourbaki school made an impressive attempt to set up mathematics as an axiomatic construction on a general and uniform basis. Indeed, it has essentially contributed to the reform of mathematics teaching at university level. The extent to which school teaching should take into account these modern tendencies is under debate. I shall return to this question later in this article.

Since the beginning of this century, axiomatic treatment has leavened very nearly all special mathematical branches. As was stressed by Einstein in different connections, the theory of relativity is based principally upon the general insight emerging from fundamental research in geometry. Without this preparatory work, the monumental interpretation of Einstein-Minkowski of the physical world in terms of four-dimensional geometry with an indefinite metric would have been inconceivable. The classical theory of Newton can also be understood and presented as a four-dimensional geometry of semi-definite metric character.

The general understanding to which fundamental geometrical research has led has also prepared the way for other unforeseen achievements in mathematics. By way of example there may be mentioned the developments of the theory of

probability during the present century (as a branch of the theory of additive set functions), information theory, econometrics, the theory of games and conclusions, which have "mathematized" research in fields where human action intervenes to an essential extent, and which have until now been considered as lying outside the scope of mathematical methods. The expansion of mathematics in directions hitherto regarded as humanistic was made possible by the very fact that mathematics is not by nature limited to what is quantitative or measurable—something now understood more clearly than ever before. It deals with different logical structures, concept complexes which have nothing to do with "magnitude" or "measurableness." One need go no further than to Euclidean geometry to comprehend this: for instance, what has incidence theory or the parallel axioms to do with the "quantitative"? Once there is correctly understood the enormous amount of generalizing and refinement contained in the mathematical mapping and correspondence concept as far as they concern the special functional connections in the world of experience or the fine arts, surprise is no longer felt that the empirical complexes accessible to such extensive conceptional idealizing as permits their being mapped into logically structured systems, and so allowing of mathematical treatment, are no longer limited to arithmetic, geometry or theories that have until now been the main concern of exact science. Many of the new fields do not lend themselves to treatment with mathematical theories formulated earlier; on the contrary, they supply empirical material and impulses for new mathematical questions and complexes of ideas. Here we see a repetition of the fascinating interplay of experience and experiment on the one hand and mathematical theory-formation on the other. This runs as a red thread through the entire evolution of exact science.

Soon after Hilbert's fundamental investigation into the foundations of elementary geometry, there were several studies which proved epoch-making on a higher level of mathematics. Hilbert's theory of linear operators and their eigenvalues and of orthogonal systems in a function space was based upon integral equations. Nevertheless, its leading idea was geometrical: Hilbert understood functional analysis to be an extension of Euclidean analytical geometry (the theory of linear and quadratic forms) into a Euclidean space with an infinite number of dimensions. The theory of Hilbert-space was given its final axiomatic form by J. von Neumann. In this form, the brilliant theory constitutes a corner-stone in the construction of modern mathematics. Its importance is not confined to "pure" mathematics. The theory of Hilbert-space assumes a leading position in the most varied connections, as for example in the modern quantum theory.

Another peculiar application of Hilbert's theory on infinite dimensional Euclidean geometry is associated with the variational problem formulated by Riemann for the purpose of solving the boundary value problem of elliptic partial differential equations of the second order. The problem is solved by the determination, within the class of all (sufficiently regular) functions with given boundary values, of the one that minimizes the Dirichlet-integral associated with

the differential equation (Dirichlet's principle). Reinterpreted in Hilbert's function-space, the problem relates to the drawing of a perpendicular, from a given point in this space, to a plane (infinite-dimensional) represented by functions with vanishing boundary values. According to Euclid, however, this perpendicular line can be characterized in two equivalent ways:

(1) The line is perpendicular to the plane.

(2) The perpendicular segment is the shortest segment joining the point and the plane.

Condition 1 implies that "the perpendicular" satisfies the differential equation, and condition 2 that the Dirichlet-integral reaches a minimum.

The elementary-geometrical minimum quality of the perpendicular in the function-space corresponds to Dirichlet's principle. This idea, as simple as it is brilliant, is one of the most fascinating indications of the enormous capacity of the structure of Euclidean geometry, apparently extremely elementary.

The above represents a brief reconsideration of some of the main features of the development of mathematics since the beginning of our century. We can learn at least *one* thing from all that has happened in mathematics during these decades of vast expansion: this period proves, if anything, the great significance, both in principle and substance, of the *Euclidean elementary system.* This theory has always stood out as the ideal of exact science. But the present century has experienced a peculiar revival of this basic mathematical theory, which has shown its enormous vitality in two different main directions: it has given impetus to the revolution in mathematics represented by the break-through of the axiomatic mode of thought, and this has had fruitful effects in the most varied branches of mathematical and logical research. But it has also, in a more substantial respect, undergone an unforeseen revival through the generalizations it has allowed in infinite-dimensional Euclidean geometry, Hilbert-space and its applications.

In view of all this, it is amazing that a group of enthusiasts supporting a radical reform of school mathematics have mustered around the slogan "Down with Euclid"! How can anything as narrow as this ever be suggested and, furthermore, by people who without doubt deserve credit for notable contributions to contemporary mathematical science?

The matter cannot be dismissed merely as a joke in bad taste, even on the assumption that the slogan has been coined mainly with a view to creating a sensation. As a symptom of a mode of thought now gaining ground in mathematics it is worth closer consideration.

In a unique way, mathematics combines two opposites: exactitude and freedom. But this *coincidentia oppositorum* also presents a danger to mathematics and mathematicians. Mathematics receives its impulses from experience and from the perception of empirical realities, but compared with other sciences, it is less rigidly tied to this starting material. On the contrary, its task is "theoretical": starting from certain phenomena characteristic of a given sphere of ex-

perience, it gradually dissociates itself from them. On a given empirical basis, it builds up a higher ideal reality. Thus mathematics, during the progress of idea-formation, leaves the empirical foundation which has given rise to the theory. This process is to a decisive extent directed by norms which can be characterized as aesthetic rather than logical: mathematics draws away from the position of real science and approaches that of creative art. Compared with artistic activity, the relative "objectivity" of mathematical creation lies mainly in its being more exact and unique.

However high mathematics may raise itself above the empirical-intuitive level—its basis historically and genetically—something of its origin still persists in even the most abstract mathematical theories. In view of the historical development of mathematics and its enormous current expansion, one can agree with ancient philosophy in the concept that mathematical theories in their abstract exactitude and beauty express a higher ideal reality in the sphere of experience from which they received their first direction-giving impulses. Each biologically and practically significant complex of the "world of experience" can be understood as a challenge to create, through a mathematical—that is, a strictly conceptional analysis, idealization, and synthesis—an image of the "structures" contained in it in its natural state.

For as long as mathematics has followed this course, equally organic and free, it has been capable of preserving its vitality. The setting up of entirely *arbitrary* axiom systems as a starting point for logical research has never led to significant results. This will scarcely be denied by any mathematician. But the awareness of this truth seems to have been dulled in the last few decades, particularly among younger mathematicians. Interest is one-sidedly directed towards formal generalizations—a rewarding field for apparent successes easily won—instead of to more difficult problems concerned with real mathematical ideas and substance. As usually happens with extremists, such a view is often associated with an over-weening arrogance towards tendencies which do not uncritically follow the modern slogans. In the wake of the revolution represented by the Bourbaki attempt—as such an imposing contribution of utmost significance—there follow a host of imitators. Here something repeats itself that has always been experienced in various connections during the development of ideas and idealogies. As before, history will decide what is of lasting value. It is the fate of imitators to fall into oblivion. But their hampering influence and the confusion of ideas they temporarily cause constitute today, perhaps, a greater danger than before, because of the expansion of mathematics and the rapidly increasing number of workers it is enlisting. In this field as well, mass phenomena increasingly assume a dominating position. Among the number of mathematical publications now flooding the world, it is not easy to find the small number which really contain some germ of a valuable idea, despite the many laudable attempts to orient research workers by means of reference papers, reviews, and monographs.

Thus research is in danger of losing contact, in some quarters, with the historic-genetic line, which alone can lead to comprehension of the vital tendencies in mathematics. Without denying the value of mathematical investigation as a high end in itself, there is good reason to point out the disadvantages of the extreme attitude, "*l'art pour l'art*," also in this branch of culture. It results in increasing uncertainty of judgment and taste.

At the same time, a distorted attitude is gaining ground with respect to what we call the "applications of mathematics." By its innermost nature, mathematical investigation in its abstract idealizations and generalizations becomes withdrawn from empirical reality, but it is accordingly given unforeseen possibilities of reacting, at a high theoretical level, within extensive areas of scientific research and culture, even those which at the initial stage of the formation of mathematical theory seem to lack any connection with the empirical basis of reality which initially inspires theoretical research. Remarkably enough, it is just the "axiomatic mode" of thought which has deepened insight into the bond uniting "theory" and "applications." Through theory-formation, mathematical structures are crystallized which, suitably reinterpreted, have applications in different fields which to the non-expert appear entirely unrelated. Along with the branching-out of science and the consequent necessary specialization one notices, in the uppermost stratum of theoretical research, a tendency directly opposed to comprehensive syntheses and greater unity.

Who could have predicted that the study of the Euclidean system, through the abstract modifications and extensions it has undergone in the multidimensional Euclidean and Riemannian general differential geometry, would one day be reinterpreted by Einstein and Minkowski in such terms that it constitutes the basis for the modern physical outlook on the world? Or that the abstract research into algebra, ultimately based on empirical and practical arithmetic, would in our time play an important role in atomic physics?

"Science seeks truth for the sake of truth," the phrase has it. This is correct. One must not then forget that the inner mainspring of mathematical research, during its really productive and vital periods, has been and is the feeling and belief, however vague, that mathematics, even in its highest theoretical speculations, has not lost contact with "empirical reality," and that it is thus called upon to take a prominent position as an important and useful component in the variety and unity of culture as a whole.

The fact that the awareness of these connections has weakened shows us the gulf which separates "pure" and "applied" mathematics today. Things were different even a few decades ago. Most leading mathematicians (Klein, Hilbert, Poincaré, Minkowski, Weyl, von Neumann) made notable contributions to the field of physics as well. The present changed situation can be explained only partially as being the necessary outcome of the further expansion and branching-out of science.

It is possible that the "applying mathematicians" will in the very near future lead the field in restoring the contacts which are lacking between theory and

practice: much more than before, "applications" depend on the support of even the most abstract theories.

In the foregoing, I have dealt mainly with certain characteristic features in the modern development of mathematical research, and have touched only briefly upon the question of mathematics teaching at school level. Discussions on school teaching rest on an uncertain basis, however, unless we can see clearly and agree upon the points which deserve primary attention in the most recent development of mathematics and its changed position in the civilized community. The reform of school teaching must be in line with these general points.

Space does not permit me to deal in detail with the consequences which may arise from the questions of principle referred to above. I restrict myself here to individual conclusions of a general nature, and hope I can add some more concrete conclusions in another connection.

It lies in the nature of things that the road to new understanding leads from the special to the general, from the concrete to the abstract. Thus it follows quite naturally that teaching has to start from concepts assumed to be known in advance, or at least easily accessible to pupils. The original context in teaching mathematics is favourable in so far that its basic material, the natural numbers in arithmetic and the simplest geometrical figures, are of such an elementary nature that pupils are largely familiar with them even before entering school. At this early stage, therefore, it is entirely needless to describe at length things that pupils already know full well. Prolonged "propaedeutics" have ill effects.

When teaching of arithmetic has advanced, more or less along old established lines, to the point where pupils intuitively comprehend the extension of the number concept by zero and negative numbers and fractions, the time is ripe for passing on systematically to symbols and to a more exact discussion of the rational number field and its basic axioms. It would be desirable to begin algebra teaching earlier than is now usual. For this purpose, the arithmetic course might be further concentrated and partly reduced. The difficulty of performing the rational operations should not be underestimated, and adequate time should be allowed for pupils to master formal mathematical manipulation. Against this, the course might be reduced by the restriction of exercises to really important and illustrative theoretical and practical examples. It is frequently stated that mathematics teaching should "approach" living life, but this demand should not be followed *irrelevantly* and *artificially*. Here, all needless subject and exercise material could be eliminated. Instead, there should be taken up questions and problems which lead from arithmetic, in a narrower sense, to algebra, to the concept of function, to geometry, and to physics. Actual instruction in algebra and geometry could then begin earlier than is now the case. Only in this way can the necessary extension of the school syllabus at higher levels be effected.

Perhaps the objection may be raised that such a concentration of arithmetic teaching is difficult because of children's limited power of understanding. In

fact, the reform, if reasonably carried out, imposes greater demands on the teachers than on the pupils. Children's capacity for understanding is easily underestimated. The present arithmetic course still includes complicated examples which cause pupils greater difficulty than do the problems which by their nature belong to a later mathematics course.

In addition, as they grow up, in some inexplicable way, young people receive impulses and accept ideas and conceptions which are "in the air." Something that might have occasioned insurmountable difficulties one or two generations ago is now more easily grasped. Human intelligence has scarcely developed in the course of history, but a shift of man's faculties has occurred. The fact that in the long run it is at all possible to orient oneself with respect to the constantly accumulating experience and knowledge, depends upon just this shift of faculties: on the one hand these are in accord with the demands of the current situation and, on the other, they give precedence to points which permit dealing with and surveying the area concerned as a whole.

The concentration and intensification of teaching arithmetic in the direction indicated above is both necessary and possible. The teaching of algebra and geometry can accordingly start without extensive preliminaries. Of course, axiomatic treatment in the strict sense of the term is not possible. Nevertheless, from the very beginning it is necessary to proceed systematically and in a disciplined manner, without getting lost, for this reason, in abstractions beyond the grasp of pupils. Exactness, consistency, and clarity do not exclude graphic description; systematic formal training need and must not degenerate into dogmatic pedantry. Under all circumstances it is the personal approach of the teacher that enlivens the teaching and arouses interest and drive in the pupils. The aim of misdirected propaedeutics, however well-intended, is achieved more efficiently *in direct association* with systematic teaching, in such a way that the presentation of the textbook, of necessity more or less schematic, is in oral teaching constantly interrupted, illustrated and augmented by digressions designed to highlight the essential points in the course itself and its relations to other relevant points.

Even the elements of algebra and geometry offer rewarding opportunities for the discussion of certain general principles of great consequence to logical thought and concept-formation as a whole. The following examples may be given:

(1) Distinction between necessary and sufficient conditions; the concept of equivalence.

(2) The logical relation between the proposition $(p \rightarrow q)$ its converse $(q \rightarrow p)$, the opposite proposition $(p \rightarrow \sim q)$, and its inverse $(q \rightarrow \sim p)$.

(3) The construction of ideal elements along the lines of the permanence principle. In fact, the extensions of the number system (from positive numbers successively to negative, rational, real and complex numbers) constitute a pattern for constructive principles, fundamental not only in mathematics but in all concept-formation.

(4) The principle of induction.

I repeat: these questions should not receive too much attention in elementary teaching. But even at this early stage, they merit consideration, in natural association with the elementary material treated, in both algebra and geometry.

(5) The significance of so-called "analysis" in solving mathematical problems.

Experience has shown it to be important that not only the beginner but also the trained mathematician be perpetually reminded of the utility of analysis. By first studying consistently the consequences which inevitably follow the assumption that there is a solution to a problem, one is led to conditions which are sufficient and at the same time indicate the path towards finding the *general* solution of the problem. On the classical model, the method is used mainly for the solution of geometrical problems, but it is equally effective in mathematical problems of all kinds—a fact frequently overlooked; an elementary example is the theory of equations at school level.

My remarks concerning elementary mathematics teaching (including the present primary and intermediate school) do not aim at a radical upheaval. There is no need for any revolutionary changes. It is important to proceed critically, with clear judgment in respect of existing advantages and disadvantages. In addition, I have pointed out the importance of combining organically the systematic and concise procedure with more general critical considerations which are to a great extent aroused by elementary mathematics. Even with pupils who do not proceed to secondary school, mathematics teaching gives the best results, from the pedagogic and practical point of view, if it is organized comprehensively in such a way as to serve as a basis for the continued study of mathematics at high school.

Uncritical attempts at "modernizing" teaching have been made by people fired by a short-sighted enthusiasm for renewal, who misunderstand the essential points in the recent development of mathematics. This is shown for instance by the proposition for a reform of algebra teaching included in the experimental courses prepared by the inter-Nordic committee. The proposal recommends that algebra should begin with a long discussion on the set concept. In an excruciatingly tedious sequence of examples, which even to children appear thoroughly trivial, time is wasted in drilling children in the use of symbols without their being given even a hint as to the possible utility of these symbols.

It is clear that from the very beginning mathematics must inevitably introduce and use abbreviating symbols (e.g. sign of equation, order sign, notations for the arithmetical basic operations or other relations, characters for quantities, for operations with sets, etc.). Well-motivated and suitably chosen notations contribute to clarification of the essentials in mathematical and logical connections. At best, they may directly influence the development of ideas (the positional system, Leibniz's notations in the infinitesimal calculus, the determinant and matrix notations, to give but a few examples).

But this is precisely why new names and notations should be introduced with

the greatest caution, and only in extremely well-motivated cases. In truth, no great capacity for invention is required for devising new arbitrary symbols. A look at current mathematical publications shows that such a tendency is increasingly being abused. It sometimes seems as if the authors wished by this means to conceal the poverty of ideas in their products.

One example of the abuse of symbols follows here. Given the equation $f(x, y) = 3x - 2y = 1$, the following problem is formulated: *Determine the set*

$$A = (x, y) \mid x \in N, \qquad y \in N, \qquad f(x, y) = 1,$$

where N means the set of all integers.

What is gained 1) in logical clarity, and 2) in time, by the employment of this complicated formulation instead of quite simply and naturally stating the problem: *Determine the integral roots* (x, y) *of equation* $f(x, y) \equiv 3x - 2y = 1$?

This is far less drastic than are the examples, filling some dozens of pages, in the inter-Nordic course referred to above. In studying mathematical textbooks, I have come across nothing poorer in ideas and more misguided than this very section.

Indeed, notations of the kind presented, are nowadays often used in scientific publications and lectures. If someone likes them, all well and good. But no special drilling is needed for the acquisition of these trivialities. Personally, in my work as teacher and scientist, I have purposely avoided unnecessary symbols, and I do not think the clarity and exactness of my presentation has suffered on this account. I do not consider it merely a matter of personal taste to refrain from using unnecessary symbols. The matter is more serious.

The use of apparently "learned" symbols in irrelevant connections tends to muddle and confuse people who lack sufficient experience and judgment. They have a superstitious respect for "knowledge," and this prevents them from using their common sense and critical judgment. Experience shows that an exaggerated and misdirected symbolism leads easily to downright superstitious and magically imbued illusions—a result which stands in sharp contrast to the actual aims of science.

A few words may be added with respect to mathematics in high school. Apart from the material dealt with along more or less conventional lines (elementary functions, the main types of conics, completion of Euclidean geometry, the course of equations, the limit concept, something about series, the rudiments of the infinitesimal calculus), the syllabus should be intensified and extended in accordance with the following points.

(1) In analytical geometry the linear transformations in the 1- and 2-dimensional case might be discussed systematically. The affine transformations between two planes are interpreted both as co-ordinate transformations and as *mappings*. In this connection, the relationship between affine and Euclidean geometry is elucidated. This can be effected wholly through discussion on quadratic forms (for two variables). The problem requires nothing not needed in any

case in connection with an equation of second degree. Euclidean metrics (the scalar product in coordinate and vector form) is then introduced, along with the orthogonal transformations and coordinate systems. This leads to the solution of the principal axis problem (eigenvalue problems).

(2) The presentation of Euclidean geometry is recapitulated and intensified, as are the basic concepts of algebra and function. At this level some glimpses of "axiomatics" are in order. In geometry, the possibility and understanding of the problematics of parallelism might be illustrated by graphic "models" from both Euclidean and non-Euclidean geometry. Analogously, the system of algebra can be analyzed by the aid of graphic examples which illustrate and generalize the elementary algebraic structures (simple cases of rings, groups, Boolean algebra, etc.).

It has been generally agreed for a long time that the "modern" concept of function, in association with the general concepts of correspondence and mapping, should be introduced at school level. Once this succeeds there is nothing to prevent discussion of the axiomatic mode of thought also in geometry, along the lines already indicated in this article. To clarify the essential points of "axiomatics," it is not necessary to carry out a thorough analysis and synthesis of an axiom system: what is required is rather an illustration of the matter with well-chosen graphic examples (this, to be sure, imposes greater demands on the teacher's ability).

(3) Similarly, the construction of the real and complex number field is recapitulated and with greater exactness than was possible in intermediate school. Here it is appropriate to discuss the limit concept.

(4) The arrangement of special courses in the elements of the infinitesimal calculus and differential equations, the theory of probability, and so on, depends on the extent to which such definite specialization can be incorporated in the high school curriculum. It must be kept in mind that a large number of pupils enter professions which do not require great quantitative knowledge of mathematics. In mathematics, the motto "non multa sed multum" should be kept in view under all circumstances.

(5) Excessive training and drill with problems should be avoided. Instead, it is of vital importance that pupils, by the solution of well-chosen and natural exercises, are given opportunities to learn the material actively and independently.

Things dangerously neglected in mathematics teaching are the applications to *physics*. Even though the present plan for the extension of school mathematics might not be fully realized, physics offers extremely rewarding material for applications. This should also be borne in mind with a view to more advanced studies.

It is not difficult to plan courses for school teaching. More difficulty is experienced in putting the plans into practice. There is more than one way of teaching well, even more of teaching badly; the worst way is that of teaching

tediously. The result of the teaching of mathematics depends in particularly large measure upon the pedagogic ability of the teachers. Unfortunately, it is scarcely possible to improve matters by methodical advice on general or special didactics. A necessary condition is that the teachers themselves have received a thorough education in mathematics and physics, and have mastered their subject far beyond the limits of the elementary material dealt with in school.

The present shortage of competent mathematics teachers is not helped by reducing university requirements: on the contrary they ought to be raised. An improvement can be obtained in quite another way: the policy should be that of making advancement possible for young teachers during their teaching career. The salaries paid in lower and higher teachers' positions should be more definitely graded. The present difficult situation in schools can be remedied only by offering increased inducements to undertaking teaching work. If there is no improvement in this respect, then any discussion on the reorganization of teaching will continue to be of no avail.

POLYNOMIALS AND POLYOMINOES

J. B. KELLY, Arizona State University

1. The associated polynomial. Let S be a finite set of lattice points (i.e. points with integral coordinates) in k-dimensional Euclidean space, E_k. There will be no loss in generality in assuming that S is contained in E_k', where E_k' is that portion of E_k in which all points have nonnegative coordinates. With the point p of S having the integral coordinates $n_1, n_2, \cdots, n_k$, we associate the monomial

$$M(p) = x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}.$$

With S itself we associate the polynomial $P(S) = \sum M(p)$, the summation being extended over all points of S. In particular, with the lattice points of the rectangular parallelopiped R, which has one vertex at the origin and lies in E_k', we associate the polynomial, $P(R)$, where

$$P(R) = \prod_{i=1}^{k} \frac{x_i^{l_i} - 1}{x_i - 1}. \tag{1}$$

Here $(l_1 - 1, l_2 - 1, \cdots, l_i - 1, \cdots, l_k - 1)$ is the point of R farthest from the origin.

Let $T_1, T_2, \cdots, T_r$ be finite sets of lattice points in E_k' and let $P(T_1)$, $P(T_2), \cdots, P(T_r)$ be their associated polynomials. We say that S is covered by $T_1, T_2, \cdots, T_r$ if every point of S is covered exactly once by suitable trans-

lations of $T_1, T_2, \cdots, T_r$ and if no point not in S is covered by these translations. This means that there exist polynomials $Q_1, Q_2, \cdots, Q_r$ in $x_1, x_2, \cdots, x_k$ with coefficients 0 or 1 such that

$$P(S) = \sum_{i=1}^{r} Q_i(x_1, \cdots, x_k)P(T_i). \tag{2}$$

(We can assume that $T_1, T_2, \cdots, T_r$ have at least one point on each coordinate axis. Then no negative exponents can occur in $Q_1, Q_2, \cdots, Q_r$.) It follows that $P(S)$ must belong to the polynomial ideal generated by $P(T_1), P(T_2), \cdots, P(T_r)$. The ring of coefficients may be any ring containing a subring isomorphic to the ring of rational integers; we shall find it convenient to employ the fields of real and complex numbers. If $(\xi_1, \xi_2, \cdots, \xi_k)$ is a point in the manifold of the ideal $(P(T_1), P(T_2), \cdots, P(T_r))$, i.e. a point with coordinates in a suitable extension of the ring of coefficients at which $P(T_1), P(T_2), \cdots, P(T_r)$ all vanish, then $P(S)$ must vanish there also. This is not, of course, a sufficient condition that $P(S)$ belong to the ideal.

To every lattice point p in E_k there corresponds a unique k-dimensional unit cube having vertices with integral coordinates, p being one of them, with no vertex having any coordinate less than the corresponding coordinate of p. (For example, in the two-dimensional case, we have a square with horizontal and vertical sides, and with p as its southwest corner.) Hence to every configuration S of lattice points there corresponds a solid region, $\bar{S}$, composed of these cubes. We set $P(\bar{S}) = P(S)$. Thus any problem involving the covering of such solid regions by other such regions may be reduced to a problem involving the corresponding configurations of lattice points. Note that (1) gives the associated polynomial of a solid rectangular parallelopiped with one vertex at the origin and sides parallel to the coordinate axes of lengths $l_1, l_2, \cdots, l_k$.

In the case $k=2$, such a "solid" region, when "rookwise" connected, was called a "polyomino" by Golomb, in his interesting paper [1] on checkerboard recreations. Here we shall use the word, "polyomino," to mean any such solid region, for any value of k. Golomb discusses problems of covering a full or deleted checkerboard with polyominoes of prescribed form. His principal tool is a "coloring" of the checkerboard. As will readily be seen, this corresponds to assigning certain values to x_1 and x_2 in our formulation.

2. Examples. In this section we shall show how (2) may be used to obtain necessary conditions for the existence of a solution of various problems involving coverings by polyominoes. Primarily we shall use the fact that $P(\bar{S})$ must vanish on the manifold of $(P(T_1), P(T_2), \cdots, P(T_r))$. It seems to be more difficult to take significant advantage of the requirement that the coefficients of $Q_1, Q_2, \cdots, Q_r$ be 0 or 1.

Example I: We begin with a well-known checkerboard problem from [1] which may easily be solved without recourse to our method of associated polynomials. Can one cover a checkerboard with one pair of opposite corners re-

moved, with 1×2 dominoes? If $\bar{S}$ is the deleted checkerboard, then

$$P(\bar{S}) = \frac{x_1^8 - 1}{x_1 - 1}\cdot\frac{x_2^8 - 1}{x_2 - 1} - 1 - x_1^7 x_2^7.$$

If $\bar{T}_1$ is the domino in its horizontal position and $\bar{T}_2$ is the domino in its vertical position then $P(\bar{T}_1) = 1 + x$, and $P(\bar{T}_2) = 1 + x_2$. From (2) we have

$$\frac{x_1^8 - 1}{x_1 - 1}\cdot\frac{x_2^8 - 1}{x_2 - 1} - 1 - x_1^7 x_2^7 = Q_1(x_1, x_2)(1 + x_1) + Q_2(x_1, x_2)(1 + x_2).$$

Setting $x_1 = -1$, $x_2 = -1$, we arrive at a contradiction, so that the desired covering is impossible. Notice that with these values of x_1 and x_2, $x_1^{n_1}x_2^{n_2}$ has one value (± 1) on all the black squares of the checkerboard in the usual coloring, and its negative (∓ 1) on the white squares. This observation relates our solution of the problem to the more elementary solution which consists simply in remarking that the proposed covering is impossible because the deleted checkerboard does not contain equal numbers of black and white squares.

The remaining examples in this section deal with the covering of rectangular k-dimensional parallelopipeds by "straight" polyominoes. By a straight polyomino we shall mean the solid region corresponding to a set of lattice points in E_k lying on a straight line parallel to one of the coordinate axes. A straight polyomino is not necessarily connected. A straight polyomino is symmetric if it is invariant under reflection in its center.

Example II: Is there some rectangular parallelopiped which can be covered by the straight symmetric polyomino formed by taking seven adjacent cubes and deleting the third and fifth? (In this, and in the subsequent examples, we agree that the polyominoes may be placed parallel to any axis.) If the polyomino is parallel to the x_i-axis, the associated polynomial for this position is $1 + x_i + x_i^3 + x_i^5 + x_i^6$. If the problem has a solution, we see, from (1) and (2), that we must have

$$\prod_{i=1}^{k} \frac{x_i^{l_i} - 1}{x_i - 1} = \sum_{i=1}^{k} Q_i(x_1, \cdots, x_k)(1 + x_i + x_i^3 + x_i^5 + x_i^6) \tag{3}$$

for some choice of the positive integers $l_1, l_2, \cdots, l_k$.

The polynomial $1 + x + x^3 + x^5 + x^6$ has a root, λ, between 0 and -1. If we put $x_1 = x_2 = \cdots = x_k = \lambda$, we obtain a contradiction from (3), inasmuch as all roots of $x_i^{l_i} - 1$ lie on the unit circle. Therefore the problem has no solution. We have made implicit use here of the theorem that a real function continuous on a closed interval assumes in that interval all values between its values at the end-points of the interval. The method of associated polynomials makes available some of the simpler theorems of analysis for the handling of problems involving lattice point configurations.

Example III: Is there a rectangular k-dimensional parallelopiped which can be covered by the straight polyomino formed by taking five consecutive cubes and deleting the middle one? Proceeding as in example II, we obtain the condition

$$
\begin{aligned}
P(\overline{S}) &= \prod_{i=1}^{k} \frac{x_i^{l_i} - 1}{x_i - 1} = \sum_{i=1}^{k} Q_i(x_1, \cdots, x_k)(1 + x_i + x_i^3 + x_i^4) \\
&= \sum_{i=1}^{k} Q_i(x_1, \cdots, x_k)(1 + x_i)^2(1 - x_i + x_i^2).
\end{aligned} \tag{4}
$$

The remainder of the argument cannot be the same as in example II because all the roots of the polynomial $1+x+x^3+x^4$ are roots of unity. Thus it is possible to select the sides l_i so that $P(S)$ vanishes on the manifold of the ideal $(1+x_1+x_1^3+x_1^4, \cdots, 1+x_k+x_k^3+x_k^4)$. But nevertheless $P(S)$ does not belong to this ideal, for any polynomial in the ideal, when expanded in powers of $1+x_1$, $1+x_2, \cdots, 1+x_k$ has no term in

$$(1 + x_1)(1 + x_2) \cdots (1 + x_k), \quad \text{whereas} \quad \frac{\partial^k P(\overline{S})}{\partial x_1 \cdots \partial x_k} \neq 0$$

when $x_1 = x_2 = \cdots = x_k = -1$ unless some $l_i = 1$. This latter case is easily excluded.

Example IV: Is there a rectangular k-dimensional parallelopiped which can be covered by the polyomino formed by taking 7 adjacent cubes and deleting the middle one?

Proceeding just as before, we obtain

$$
\begin{aligned}
P(\overline{S}) &= \prod_{i=1}^{k} \frac{x_i^{l_i} - 1}{x_i - 1} = \sum_{i=1}^{k} Q_i(x_1, \cdots, x_k)(1 + x_i + x_i^2 + x_i^4 + x_i^5 + x_i^6) \\
&= \sum_{i=1}^{k} Q_i(x_1, \cdots, x_k) \frac{(x_i^4 + 1)(x_i^3 - 1)}{x_i - 1}.
\end{aligned} \tag{5}
$$

Again, the roots of $1+x+x^2+x^4+x^5+x^6$ are all roots of unity. In this case, however, $P(\overline{S})$ will belong to the ideal $(1+x_1+x_1^2+x_1^4+x_1^5+x_1^6, \cdots, 1+x_k+x_k^2+x_k^4+x_k^5+x_k^6)$ if the integers l_i are divisible by 24. This follows from the fact that $x^{24}-1$ is divisible by $(x^3-1)(x^4+1)$. To handle the problem it is necessary then to make use of the condition that the coefficients of the polynomials $Q_i(x_1, x_2, \cdots, x_k)$ be 0 or 1, or possibly, of the weaker condition, that they be nonnegative. We have had no success with this; we can say only that no solution exists when $k=2$, a fact established by trial and error. The case $k>2$ is open.

Example V: The preceding examples may lead one to suspect that any straight, symmetric polyomino which cannot cover any segment, cannot cover any rectangular parallelopiped. A polyomino formed by taking 6 adjacent

cubes and removing the second and fifth obviously cannot cover any segment, but such a polyomino can cover a 7×12 rectangle. This is shown in Figure 1. There the polyominoes are numbered from 1 to 21 and a square numbered a, $1\leqq a\leqq 21$, is covered by the polyomino numbered a.

15	8	16	8	8	13	8	13	13	19	13	21
5	14	5	5	17	5	12	18	12	12	20	12
15	4	16	4	4	11	4	11	11	19	11	21
15	14	16	7	17	7	7	18	7	19	20	21
3	14	3	3	17	3	10	18	10	10	20	10
15	2	16	2	2	9	2	9	9	19	9	21
1	14	1	1	17	1	6	18	6	6	20	6

FIG. 1

3. The box problem. Most of the results in the preceding section were of a negative character. In this section we shall discuss what is perhaps the simplest problem of polyomino coverings, obtain a necessary condition for its solvability by means of the method of associated polynomials and then show that this necessary condition, together with an auxiliary condition, is sufficiently strong to guarantee the existence of a solution.

The problem, which we have called the box problem, is the following: Under what circumstances is it possible to stack k-dimensional "boxes" with integral sides $b_1, b_2, \cdots, b_k$ in a k-dimensional "room" with sides $r_1, r_2, \cdots, r_k$ so that the room is completely filled? Clearly, the volume of one of the boxes must divide the volume of the room and each of the numbers $r_1, r_2, \cdots, r_k$ must be a linear combination of $b_1, b_2, \cdots, b_k$, with nonnegative integral coefficients. We prove a demonstrably stronger necessary condition:

(A) *If an arbitrary integer h divides t_h of the integers $b_1, b_2, \cdots, b_k$, it must divide at least t_h of the integers $r_1, r_2, \cdots, r_k$.*

It follows from (A) that, for example, a 30×30 square cannot be covered by 4×9 rectangles even though 30×30 is divisible by 4×9 and $30=2\cdot 9+3\cdot 4$.

Proof of (A). From (1) and (2) we have

$$\prod_{i=1}^{k}\frac{x_i^{r_i}-1}{x_i-1}=\sum_{\sigma}Q_\sigma(x_i,\cdots,x_k)\prod_{i=1}^{k}\frac{x_{\sigma(i)}^{b_i}-1}{x_{\sigma(i)}-1},\tag{6}$$

the summation being extended over all permutations σ of the integers $1, 2, \cdots, k$. Each permutation corresponds to a different way of stacking the boxes.

Suppose that only q of the integers $r_1, r_2, \cdots, r_k$ are divisible by h, where $q<t_h$. Then $k-q$ of the integers $r_1, r_2, \cdots, r_k$ are not divisible by h. We may assume that these are $r_1, r_2, \cdots, r_{k-q}$. In (6), let $x_1=x_2=\cdots=x_{k-q}=\omega$ where ω is a primitive hth root of unity. In each product on the right side of (6) there is at least one factor which vanishes, since $k-q+t_h>k-q+q=k$. Thus the right side of (6) vanishes identically in $x_{k-q+1}, \cdots, x_k$, whereas the left side does not.

Condition (A) is clearly not sufficient. For example, it is impossible to fill a $48\times 48\times 1$ room with $2\times 3\times 4$ boxes, although condition (A) is satisfied. What is needed is an additional condition which ensures that $r_1, r_2, \cdots, r_k$ may be expressed as linear combinations, with positive integral coefficients, of $b_1, b_2, \cdots, b_k$. Such a condition is

(B) $r_1, r_2, \cdots, r_k$ *are sufficiently large.*

That is, there exists a positive integer, N, such that if $r_i>N$, $i=1, 2, \cdots, k$, and the set $\{r_1, r_2, \cdots, r_k\}$ satisfies condition (A), then the box problem has a solution. Here N depends upon $b_1, b_2, \cdots, b_k$.

We prove that conditions (A) and (B) are sufficient for the existence of a solution of the box problem. Our plan is to split the room into smaller parallelopipeds, each of whose sides is divisible by a different number in the set $b_1, b_2, \cdots, b_k$. These smaller parallelopipeds are then obviously coverable by boxes of sides $b_1, b_2, \cdots, b_k$; hence the room is also.

Two lemmas, both well-known results, are required.

LEMMA I. *Let $a_1, a_2, \cdots, a_n$ be any set of positive integers, and let δ be their greatest common divisor. Any sufficiently large integer which is divisible by δ may be expressed as a linear combination of $a_1, a_2, \cdots, a_n$ with positive integral coefficients.*

An account of recent work based upon this lemma may be found in [2]. It was known to Frobenius, and may have been noticed by earlier mathematicians.

LEMMA II. *Given k objects, each of which possesses one or more of the attributes $P_1, P_2, \cdots, P_k$. For any set of j attributes, let there exist j objects each possessing at least one attribute of the set. Then it is possible to pair each object with one of its attributes in such a way that no two objects are paired with the same attribute.*

This lemma is due to P. Hall [3].

Let us order all nonempty subsets of $b_1, b_2, \cdots, b_k$ by means of an index j, running from 1 to 2^k-1. Let $b_{j_1}, b_{j_2}, \cdots, b_{jm_j}$ be the elements of the jth subset. Let δ_j be the greatest common divisor of $b_{j_1}, b_{j_2}, \cdots, b_{jm_j}$. Condition (A) implies that δ_j divides at least m_j of the integers $r_1, r_2, \cdots, r_k$.

Suppose that $j_1, j_2, \cdots, j_{\alpha_i}$ are the indices of those greatest common divisors which divide r_i. Condition (A) gives $\alpha_i \geqq 1$. Then we may write

$$r_i = \sum A^{(i)}_{j_1 l_1, j_2 l_2, \cdots, j_{\alpha(i)} l_{\alpha(i)}}, \tag{7}$$

where the summation extends over all possible sets of values of $l_1, l_2, \cdots, l_{\alpha_i}$ such that

$$1 \leqq l_1 \leqq m_{j_1},\ 1 \leqq l_2 \leqq m_{j_2}, \cdots, 1 \leqq l_{\alpha_i} \leqq m_{\alpha_i},$$

and where the positive integer $A^{(i)}_{j_1l_1,j_2l_2,\cdots,j_{\alpha_i}l_{\alpha_i}}$ is divisible by each of the integers $b_{j_1l_1}, b_{j_2l_2}, \cdots, b_{j_{\alpha_i}l_{\alpha_i}}$. This result follows at once from the fact that the positive integers form a distributive lattice under the operations $\cap$ = least common multiple and $\cup$ = greatest common divisor. One infers that r_i is divisible by the greatest common divisor of all the least common multiples one can form by taking one b from each set associated with the α_i δ's dividing r_i. If we apply Lemma I, identifying the integers $a_1, a_2, \cdots, a_n$ with these least common multiples, we see that the integers $A^{(i)}_{j_1l_1,j_2l_2,\cdots,j_\alpha l_\alpha}$ may be chosen to be positive if r_i is sufficiently large.

We now divide our k-dimensional room into smaller parallelopipeds by splitting the sides as indicated by (7). The sides of a representative smaller parallelopiped will be

$$A^{(1)}_{j_1l_1,j_2l_2,\cdots,j_{\alpha_1}l_{\alpha_1}} \cdots A^{(k)}_{j_1'l_1'\cdots j_{\alpha_k}'l_{\alpha_k}'}. \tag{8}$$

We apply Lemma II, the k objects being the sides of the smaller parallelopiped and the k attributes being divisibility by $b_1, b_2, \cdots, b_k$. We show that the hypothesis of Lemma II is fulfilled. Let δ_j be the greatest common divisor of some set of m_j b's. Then δ_j will divide at least m_j of the integers $r_1, r_2, \cdots, r_k$; by our construction, therefore, at least m_j of the positive integers (8) are divisible by at least one member of the given set of b's. From Lemma II we conclude that it is possible to pair each side of the smaller parallelopiped with a distinct member of the set $b_1, b_2, \cdots, b_k$ which will divide it. Thus this parallelopiped may be covered with boxes of sides $b_1, b_2, \cdots, b_k$. The larger room with sides $r_1, r_2, \cdots, r_k$ may then also be so covered.

N. G. de Bruijn, in a problem published in the Hungarian journal, *Mathematikai Lapok*, around 1960, dealt with an interesting aspect of the box-problem. He showed that if the box problem has a solution and if b_1 divides b_2, b_2 divides b_3, etc., then the boxes may all be given the same orientation. Moreover, if every room covered by boxes of dimensions $b_1 \leqq b_2 \leqq b_3 \leqq \cdots \leqq b_n$ may be covered by boxes with the same orientation, then b_1 divides b_2, b_2 divides b_3, etc. In the course of his proof, de Bruijn established the necessity portion of our box problem for boxes with dimensions $1 \times 1 \times 1 \times \cdots \times m$. His method was similar to ours and may be extended to the more general case.

Conclusion. Certain mathematical games involving "jumping" may be discussed by means of associated polynomials. One allows the associated polynomials to have coefficients $-1, 0, 1$ in such a case. In treating questions involving multiple covering of lattice points, the sole restriction that one need place on $Q_1, Q_2, \cdots, Q_r$ is that they have nonnegative integral coefficients.

Interesting problems arise when one considers infinite sets of lattice points.

Here the associated polynomial becomes an associated formal power series. If one substitutes real or complex numbers for the indeterminates, convergence difficulties present themselves, particularly if the configuration extends from $-\infty$ to $+\infty$ in some direction.

Fundamentally, the method of attack in this paper goes back to Descartes. The associated polynomial is merely the "coordinate" of the polyomino.

References

1. S. Golomb, Checkerboards and polyominoes, this MONTHLY, 61 (1954) 675–682.

2. A. Brauer and J. Shockley, On a problem of Frobenius, J. R. Angew. Math., 211 (1962) 215–220.

3. P. Hall, On representatives of subsets, J. London Math. Soc., 10 (part I) (1935) 26–30.

ON A DIOPHANTINE EQUATION WITH NO NONTRIVIAL INTEGRAL SOLUTION

RAPHAEL FINKELSTEIN, University of Arizona

In 1940 the following problem was proposed for solution in the MONTHLY [5]: *Does the equation*

$$m^3 + 3m^2 + 2m = 2n^3 + 3n^2 + n \tag{1}$$

admit positive integral solutions in m and n other than $m=n=1$?

Since then no solution has been received for this problem, though Makowski showed in 1958 [3] that equation (1) has no integral solutions if m and n are integers greater than 1 and less than 3164, and $|2m-n| \leqq 100$. This result is hardly surprising, for in this paper we shall prove the following

THEOREM. *The only solutions of equation* (1) *in nonnegative integers are* $[m, n] = [0, 0]$ *and* $[1, 1]$.

Proof. If in equation (1) we put $x=2n+1$, $y=-(m+1)$, it becomes

$$x^3 + 4y^3 = x + 4y. \tag{2}$$

We shall prove several lemmas which together imply that the only nonzero integral solutions of (2) subject to the conditions

$$x \text{ positive and odd, } y \text{ negative,} \tag{3}$$

are $[x, y]=[1, -1]$ and $[3, -2]$, which will prove the result.

In the course of our investigations we shall require the following theorems, which we state here without proof.

THEOREM A (LAGRANGE). *If* $m>1$, *the only integral solution of the Diophantine equation* $x^3+2^my^3=1$ *is* $[x, y]=[1, 0]$.

THEOREM B (NAGELL [4]). *The Diophantine equation* $Ax^3+By^3=C$ ($C=1$ *or* 3; $3\nmid AB$ *if* $C=3$; $A>B$; A, B, C *positive integers*) *has at most one solution in nonzero integers* $[x, y]$. *There is the unique exception for the equation* $2x^3+y^3=3$, *which has exactly the two integral solutions* $[x, y]=[1, 1]$ *and* $[4, -5]$.

THEOREM C (HEMER [2, 69–70]). *Let* ϵ *be a unit in a cubic ring and let the odd prime* p *be a divisor of* $N(\epsilon'+\epsilon'')$. *Suppose further that* $\epsilon^m=a_m\epsilon^2+b_m\epsilon+c_m$ *is the least power of* ϵ *with* $m>0$ *for which* $a_m\equiv b_m\equiv 0 \pmod p$. *Then* $\epsilon^n=u+v\epsilon$ *has no solution for even* n *except* $n=0$ *if* $a_m\not\equiv 0 \pmod{p^2}$, *and no solution for odd* n *except* $n=1$ *if* $c_{m+2}\not\equiv 0 \pmod{p^2}$.

We now return to our problem and prove first

LEMMA 1. *All the integral solutions of* (2), *subject to conditions* (3), *correspond to the integral solutions of the equations*

$$a^3 + 4b^3 = \pm 1, \pm 3, \pm 5, \text{ and } \pm 15. \tag{4}$$

Proof. Let $[x, y]$ be any integral solution of (2), subject to conditions (3). We put $(x, y)=d$, where d is the greatest common divisor of x and y. Then, on substituting $x=ad$, $y=bd$ in (2) and simplifying, we have

$$d^2(a^3 + 4b^3) = a + 4b, \tag{5}$$

where

$$(a, b) = 1, \ a \text{ is positive and odd, } b \text{ is negative and } d \text{ is odd.} \tag{6}$$

We set

$$a^3 + 4b^3 = c. \tag{7}$$

Then $(a, c)=1$, and

$$a = cd^2 - 4b. \tag{8}$$

Now we substitute (8) for a in (7) and expand, obtaining

$$c = c^3d^6 - 12c^2d^4b + 48cd^2b^2 - 60b^3. \tag{9}$$

But from (7)

$$-60b^3 = -15c + 15a^3, \tag{10}$$

so we can substitute (10) in (9) to get

$$-15a^3 = c^3d^6 - 12c^2d^4b + 48cd^2b^2 - 16c.$$

Hence, $c\mid 15a^3$, and thus $c\mid 15$, because $(a^3, c)=1$. Thus $c=\pm 1$, ± 3, ± 5 or ± 15, and solving (2) in integers, subject to conditions (3), is equivalent to solving the set of equations

$$a^3 + 4b^3 = \pm 1, \pm 3, \pm 5 \text{ and } \pm 15,$$

subject to conditions (6). This proves Lemma 1.

LEMMA 2. *The only nonzero integral solution of the equations* $a^3+4b^3=\pm 1$ *and* ± 3 *fulfilling conditions* (6) *is* $[a, b]=[1, -1]$.

Proof. First, we note that $[a, b]$ is a solution of $a^3+4b^3=-c$ if and only if $[-a, -b]$ is a solution of $a^3+4b^3=c$, since the latter equation may be written as $(-a)^3+4(-b)^3=-c$. Next, Theorem A shows that the only integral solution of $a^3+4b^3=1$ is $[a, b]=[1, 0]$, and hence the only integral solution of $a^3+4b^3=-1$ is $[a, b]=[-1, 0]$. Finally, by Theorem B, the only integral solution of $a^3+4b^3=3$ is $[a, b]=[-1, 1]$, and hence the only integral solution of $a^3+4b^3=-3$ is $[a, b]=[1, -1]$. The only one of these solutions satisfying conditions (6), however, is $[a, b]=[1, -1]$.

LEMMA 3. *The equations* $a^3+4b^3=\pm 15$ *are both impossible in rational integers* $[a, b]$.

Proof. It suffices, by the remark just following the statement of Lemma 2, to prove the impossibility of the equation $a^3+4b^3=15$. The method to be used requires a discussion of the cubic field $K(\theta)$, where $\theta^3=2$. The necessary information for our problem is provided in articles by Nagell [4] and Dedekind [1], and is as follows:

i. The integers of $K(\theta)$ take the form $\alpha=A+B\theta+C\theta^2$, where A, B and C are rational integers, i.e., $(1, \theta, \theta^2)$ is an integral basis for $K(\theta)$.

ii. The ring of integers of $K(\theta)$ is a unique factorization domain.

iii. By Dirichlet's theorem on units, there is only one fundamental unit of the field, which we designate by ϵ_0. From the table on page 120 of [1] ϵ_0, where $0<\epsilon_0<1$, is given by

$$\epsilon_0 = -1+\theta.$$

All units of the field are given by $\pm\epsilon_0^m$, where m is any rational integer. Any such power of ϵ_0 is of the form $u+v\theta+w\theta^2$, where u, v and w are rational integers. The product uvw is zero only when $m=0$ and $m=1$. All units of norm 1 of $K(\theta)$ are given by ϵ_0^m.

iv. Since $x^3-2\equiv(x-2)^3 \pmod 3$, the general theory of algebraic numbers shows that 3 is a perfect cube in $K(\theta)$, apart from unit factors. We have, in fact,

$$3 = (\theta+1)^3(\theta-1),$$

and 3 is the cube of a prime of norm 3 times a unit factor, since $N(\theta-1)=1$. Hence, there is only one equivalence class of associated primes of norm 3 in $K(\theta)$, as any integer of norm 3 in $K(\theta)$ must divide 3, and, apart from unit factors, there is only one such integer.

v. Since 5 is a rational prime of the form $3r+2$, 5 does not divide the discriminant of the field, -108, and $-108\equiv 2 \pmod 5$ while 2 is a nonquadratic residue modulo 5, the integer 5 splits into two primes in $K(\theta)$, one of norm 5

and one of norm 25. We have, in fact,

$$5 = (\theta^2 + 1)(-\theta^2 + 2\theta + 1), \tag{11}$$

and it is readily verified that each integer on the right-hand side of (11) is a prime of $K(\theta)$. Since the norm of $\alpha = A + B\theta + C\theta^2$ is given by

$$N(\alpha) = A^3 + 2B^3 + 4C^3 - 6ABC,$$

the first of the factors of 5 shown in (11) has norm 5, the second norm 25. Hence, there is only one equivalence class of associated primes of norm 5 in $K(\theta)$, as any integer in $K(\theta)$ which has norm 5 must divide 5, and, apart from unit factors, there is only one such integer.

We consider now the integer $-1+2\theta$. Its norm is 15, and any other integer of norm 15 must be associated with $-1+2\theta$, as all primes of norm 5 in $K(\theta)$ are associated, as are all primes of norm 3. We seek all integers of norm 15 in $K(\theta)$ which are of the form $a+b\theta^2$.

Let $\epsilon_0^m = u_m + v_m\theta + w_m\theta^2$ be a unit of $K(\theta)$ of norm 1. We require the coefficient of θ in $(-1+2\theta)(u_m+v_m\theta+w_m\theta^2)$ to be zero. This yields

$$-v_m + 2u_m = 0. \tag{12}$$

We shall show that (12) is impossible, which will imply Lemma 3. We now set

$$x_m + y_m\theta + z_m\theta^2 = (-1 + 2\theta)\epsilon_0^m. \tag{13}$$

Since

$$\begin{gathered} \epsilon_0^3 + 3\epsilon_0^2 + 3\epsilon_0 - 1 = 0, \text{ i.e.,} \\ \epsilon_0^{m+3} + 3\epsilon_0^{m+2} + 3\epsilon_0^{m+1} - \epsilon_0^m = 0, \end{gathered} \tag{14}$$

and $(1, \theta, \theta^2)$ is an integral basis for $K(\theta)$, it follows at once that

$$y_{m+3} + 3(y_{m+2} + y_{m+1}) - y_m = 0. \tag{15}$$

From (13), we have $x_m + y_m\theta + z_m\theta^2 = (-1+2\theta)(u_m+v_m\theta+w_m\theta^2)$, and thus

$$y_m = - v_m + 2u_m = 0 \tag{16}$$

by (12).

Since $y_m = 0$, the congruence $y_m \equiv 0 \pmod t$ must be solvable for every modulus t. We shall prove Lemma 3 by showing that the congruence $y_m \equiv 0 \pmod 5$ is impossible.

To this end we note that $\epsilon_0^8 = 1 + 100\theta - 80\theta^2$, i.e., $\epsilon_0^8 \equiv 1 \pmod 5$. Hence, we must have $y_{m+8} \equiv y_m \pmod 5$. Thus, to prove the impossibility of (12), it is sufficient to show that

$$y_i \ (i = 0, 1, \cdots, 7) \not\equiv 0 \pmod 5.$$

By actual computation, using (15) and (16), we find the values of y_0 to y_7,

modulo 5, to be

$$2, 2, 4, 4, 3, 3, 1 \text{ and } 1,$$

and none of these are zero. This proves Lemma 3.

LEMMA 4. *The only integral solutions of the equations* $a^3+4b^3=\pm 5$ *are* $[a, b]=[1, 1]$, $[-3, 2]$, $[-1, -1]$ *and* $[3, -2]$. *Of these, only* $[3, -2]$ *satisfies conditions* (6).

Proof. As before, it suffices to show that the only integral solutions of the equation $a^3+4b^3=5$ are $[a, b]=[1, 1]$ and $[-3, 2]$.

We seek all the primes of $K(\theta)$ of norm 5 which are of the form $a+b\theta^2$. Since all primes of norm 5 in $K(\theta)$ are associated, any such prime must be an associate of $1+\theta^2$, which is an integer of norm 5. Let $\bar{\epsilon}=X+Y\theta+Z\theta^2$ be a unit of $K(\theta)$ of norm 1. We require the coefficient of θ in

$$(1 + \theta^2)(X + Y\theta + Z\theta^2)$$

to be zero. This yields.

$$Y = -2Z. \tag{17}$$

Since

$$N(\bar{\epsilon}) = X^3 + 2Y^3 + 4Z^3 - 6XYZ = 1, \tag{18}$$

we can substitute (17) into (18) to get

$$X^3 + 12XZ^2 - 12Z^3 = 1. \tag{19}$$

In (19) let $X=T+Z$. Then (19) becomes

$$T^3 + 3T^2Z + 15TZ^2 + Z^3 = 1. \tag{20}$$

We shall now show that the only integral solutions of (20) are $[T, Z]=[1, 0]$ and $[0, 1]$. From this will follow that the only integral solutions of (19) are $[X, Z]=[1, 0]$ and $[1, 1]$, which will show that the only units of $K(\theta)$ of norm 1 which yield numbers of norm 5 of the form $a+b\theta^2$ when multiplied by $1+\theta^2$ are 1 and $1-2\theta+\theta^2$, which will prove that the only such numbers are $1+\theta^2$ and $-3+2\theta^2$, and this will prove Lemma 4.

We consider now the ring $R(1, \epsilon, \epsilon^2)$, designated by $R(\epsilon)$, defined by

$$\epsilon^3 - 3\epsilon^2 + 15\epsilon - 1 = 0. \tag{21}$$

Since $R(\epsilon)$ has discriminant $-10{,}800$, there is one fundamental unit of the ring, and the group of all units of $R(\epsilon)$ consists of the integral powers of this unit and their negatives. Clearly, ϵ is a unit of $R(\epsilon)$. We show that it may be taken as the fundamental unit of $R(\epsilon)$. If we solve (21) by Cardan's method we obtain $\epsilon=1-2\theta+\theta^2$, i.e., $\epsilon=(-1+\theta)^2$. Now the integers of $K(\theta)$ form a maximal cubic lattice. Hence, since $-1+\theta$ is the fundamental unit of the ring of integers of $K(\theta)$, it is not a power of another cubic unit. Thus, since $-1+\theta \notin R(\epsilon)$, ϵ is not

a power of another unit of $R(\epsilon)$, and hence it may be taken as the fundamental unit of $R(\epsilon)$.

We now apply Theorem C. We have $N(\epsilon'+\epsilon'')=N(3-\epsilon)=N(2+2\theta-\theta^2)=44$. Hence, $p=11$. We must find the smallest power of ϵ, ϵ^m, such that if $m>0$ and $\epsilon^m=A_m\epsilon^2+B_m\epsilon+C_m$, $A_m\equiv B_m\equiv 0 \pmod{11}$.

To this end, we note that $\epsilon^{m+1}=\epsilon(A_m\epsilon^2+B_m\epsilon+C_m)$, which implies that

$$(22)\qquad C_{m+1}=A_m,\quad B_{m+1}=C_m-15A_m,\quad \text{and}\quad A_{m+1}=3A_m+B_m.$$

Also

$$\epsilon^{10}=-48{,}267\epsilon^2+2728\epsilon+36,\quad \text{i.e.,}\quad \epsilon^{10}\equiv(\epsilon^2+3) \pmod{11}.$$

This implies that $\epsilon^{20}\equiv 1 \pmod{11}$, so that

$$A_{m+20}\equiv A_m \pmod{11},\qquad B_{m+20}\equiv B_m \pmod{11}.$$

Using (21) and (22), we have made a table of values of A_m and B_m to the moduli 11 and 121 for $m=0, 1, \cdots, 21$. (See Fig. 1.)

m	A_m (mod 11)	B_m (mod 11)	A_m (mod 121)
0	0	0	0
1	0	1	0
2	1	0	1
3	3	−4	3
4	−6	0	−6
5	4	5	59
6	−5	0	28
7	−4	2	40
8	1	0	1
9	3	3	36
10	1	0	12
11	3	−1	−19
12	−3	0	41
13	2	4	57
14	−1	0	21
15	−3	−5	−25
16	−3	0	30
17	2	−2	2
18	4	0	15
19	1	−3	45
20	0	0	33
21	0	1	44

The table shows that $m=20$. Since $A_{20}\equiv 33 \pmod{121}$ and $C_{22}=A_{21}\equiv 44 \pmod{121}$, $[T, Z]=[1, 0]$ and $[0, 1]$ by Theorem C. Hence, the only integral solutions of $a^3+4b^3=5$ are $[a, b]=[1, 1]$ and $[-3, 2]$, and thus the only integral solutions of $a^3+4b^3=-5$ are $[a, b]=[-1, -1]$ and $[3, -2]$. Of all these solutions, the only one satisfying conditions (6) is $[a, b]=[3, -2]$.

LEMMA 5. *For integral solutions* $[x, y]$ *of equation* (2), *subject to conditions* (3), $d=1$.

Proof. Lemmas 2 to 4 show that the only solutions of equations (4) satisfying conditions (6) are $[a, b]=[1, -1]$ and $[3, -2]$. Thus, from (5) we get

$$-3d^2 = -3, \quad -5d^2 = -5, \quad \text{i.e.,} \quad d = 1.$$

Thus, the only integral solutions of equation (2), subject to conditions (3), are $[x, y]=[1, -1]$ and $[3, -2]$, and thus the only solutions of (1) in nonnegative integers are $[m, n]=[0, 0]$ and $[1, 1]$. The proof is complete.

Note: The problem which is solved in this paper is mentioned on page 96 of *Tomorrow's Math* by C. S. Ogilvy.

References

1. R. Dedekind, Über die Anzahl der Idealklassen in reinen kubischen Zahlkörpern, Journ. für Math., 121 (1900) 40–123.
2. O. Hemer, Notes on the Diophantine Equation $y^2-k=x^3$, Ark. Mat., 3 (1954) 67–77.
3. A. Makowski, Comment on Problem 3951, this MONTHLY, 65 (1958) 371.
4. T. Nagell, Solution complète de quelques équations cubiques à deux indéterminées, J. Math. Pures Appl., Serie IX, 4 (1925) 209–270.
5. Problem 3951, this MONTHLY, 47 (1940) 182.

AREAS AND VOLUMES WITHOUT LIMIT PROCESSES

D. E. RICHMOND, Williams College

1. Introduction. It is universally believed that areas under polynomial graphs can be found only by some sort of limit process. This is not the case!

We shall assume as usual that the area of a plane region bounded by a simple closed curve is a positive number, that congruent regions have the same area, and that the area of the union of two nonoverlapping regions is the sum of their areas. Finally, we assume that the area of a rectangle is the product of the length of its base by its altitude.

We begin with a simple case.

2. The area under a parabola. Let $F(x)$ be the area under the parabola $y=x^2$ above the interval $[0, x]$ and $F(x')$ the area above $[0, x']$ $(x'>x)$. Then the area above $[x, x']$ is $F(x')-F(x)$ (additivity). This area is greater than that of the rectangle of base $x'-x$ and altitude x^2, and less than the area of the rectangle of base $x'-x$ and altitude x'^2 (additivity and positivity). Hence,

$$(x' - x)x^2 < F(x') - F(x) < (x' - x)x'^2. \tag{1}$$

This must be true for all $0 \leqq x < x'$.

It is very easy to find a function F such that the double inequality (1) is satisfied for all $0 \leqq x < x'$. $F(x)$ is clearly less than half the area of the rectangle of base x and altitude x^2, so that $F(x) = x^3/3$ (and $F(x') = x'^3/3$) is a good guess.

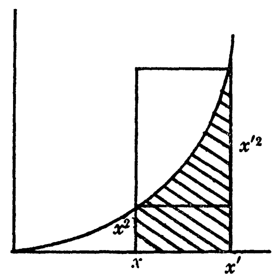


FIG. 1

In fact,

$$\frac{x'^3}{3} - \frac{x^3}{3} = (x' - x)\left(\frac{x^2 + xx' + x'^2}{3}\right)$$

is less than $(x'-x)x'^2$ and greater than $(x'-x)x^2$.

If desired, one can start with the easily derived inequality

$$(x' - x)3x^2 < x'^3 - x^3 < (x' - x)3x'^2$$

and divide by 3.

$F(x) = x^3/3$ is of course the correct solution. To justify it we must show that if there exists an area function F for which $F(0) = 0$ and

$$(x' - x)x^2 < F(x') - F(x) < (x' - x)x'^3$$

for all $0 \leqq x < x'$, then this function is unique. We need not prove the *existence* of such a function, since a bird in the hand exists.

The essence of the method is to note that for a function f which increases on an interval $[a, b]$, an area function $F(x)$ must satisfy the double inequality

$$(2) \qquad (x' - x)f(x) < F(x') - F(x) < (x' - x)f(x')$$

for all $x < x'$ on the interval. For elementary functions, F can be found easily without dividing by $x' - x$ and passing to the limit as $x' \to x$ as is customary. The exhibition of such a function F makes it unnecessary to prove its existence. If $F(a) = 0$, which is easy to arrange, the theorem to be proved in the next section shows that $F(x)$ is the only possible answer for the required area above $[a, x]$.

Since we can treat decreasing functions by reversing the inequality signs in (2), this simple method is sufficient to handle elementary functions which are piecewise monotone.

3. The uniqueness proof.

THEOREM 1. *Let f be a nonnegative strictly increasing function on the interval $a \leqq x \leqq b$. Let $F(x)$ be the area below the graph of $y=f(x)$ and above the interval $[a, x]$. If $F(a)=0$ and*

$$(x' - x)f(x) < F(x') - F(x) < (x' - x)f(x') \tag{2}$$

for all $a \leqq x < x' \leqq b$, F is uniquely defined.

Proof. Assume that there exists a different function $\bar{F}$ satisfying (2) with $\bar{F}(a)=0$ so that for some $c(a<c\leqq b)$, $\bar{F}(c)\neq F(c)$.

Let

$$x_k = a + \frac{k(c-a)}{n},$$

n a positive integer and $k=0, 1, 2, \cdots, n$. Then

$$\frac{c-a}{n} f(x_{k-1}) < F(x_k) - F(x_{k-1}) < \frac{c-a}{n} f(x_k) \qquad (k = 1, 2, \cdots, n).$$

Summing from $k=1$ to $k=n$,

$$\frac{c-a}{n}\sum_{0}^{n-1} f(x_k) < F(c) < \frac{c-a}{n}\sum_{1}^{n} f(x_k).$$

Similarly,

$$\frac{c-a}{n}\sum_{0}^{n-1} f(x_k) < \bar{F}(c) < \frac{c-a}{n}\sum_{1}^{n} f(x_k).$$

Hence, $|\bar{F}(c)-F(c)| < [(c-a)/n][f(c)-f(a)]$ and

$$n < \frac{(c-a)[f(c)-f(a)]}{|\bar{F}(c) - F(c)|}.$$

Since the Archimedean axiom assures us that there exists a positive integer n for which

$$n|\bar{F}(c) - F(c)| > (c-a)[f(c) - f(a)],$$

the assumption that $\bar{F}(c)\neq F(c)$ leads to a contradiction and must be rejected.

The proof is easily modified to treat decreasing functions where the inequality signs in (2) are reversed.

It will be observed that this proof does not require that f be continuous. (I am indebted to Professor Robert H. Breusch for suggesting an improvement of a previous proof which did assume the continuity of f. See also Apostol, *Calculus*, Vol. I, pp. 7–8.) It therefore appears that no use has been made of the limit concept. Nor has the Cantor continuity postulate for the reals been used.

4. Areas under polynomial graphs. The double inequality

$$(2) \qquad (x' - x)f(x) < F(x') - F(x) < (x' - x)f(x') \qquad (0 \leqq x < x')$$

is easily solved if $f(x) = cx^n$, $c > 0$ and n a positive integer. Since

$$x'^{n+1} - x^{n+1} = (x' - x)(x^n + x^{n-1}x' + \cdots + x'^n)$$

is between $(x' - x)(n+1)x^n$ and $(x' - x)(n+1)x'^n$ we readily find

$$F(x) = \frac{cx^{n+1}}{n+1}.$$

Of course, $F(0) = 0$.

To handle polynomial functions, we need two simple theorems.

THEOREM 2. *If f and g are nonnegative strictly increasing functions on $x \geqq 0$ and F and G are their respective area functions, then $F+G$ is the area function for $f+g$.*

Proof. We add

$$(x' - x)f(x) < F(x') - F(x) < (x' - x)f(x')$$

and

$$(x' - x)g(x) < G(x') - G(x) < (x' - x)g(x')$$

and obtain

$$(x' - x)[f(x) + g(x)] < [F(x') + G(x')] - [F(x) + G(x)] < (x' - x)[f(x') + g(x')].$$

Note that $(F+G)(0) = F(0) + G(0) = 0$.

Repeated use of this theorem gives the areas under polynomial graphs over any interval $[a, b]$, $a \geqq 0$, provided that all the coefficients are positive.

For polynomials some of whose coefficients are negative, we may write the polynomial function as the difference of two increasing polynomial functions, p and q. ($f = p - q$.) We assume that we are dealing with an interval for which $p(x) > q(x)$. The graph of $f(x)$ consists of a finite number of pieces over which f increases or decreases. For definiteness, assume that f increases on the interval $[a, b]$.

THEOREM 3. *Let p and q be nonnegative, strictly increasing functions on $[a, b]$ and let P and Q be the corresponding area functions. If $f = p - q$ is positive and increasing (decreasing) on $[a, b]$, then $F = P - Q$ is the area function for f.*

Geometric Proof. The area above $[x, x'](a \leqq x < x' \leqq b)$ and below $y = p(u)$ $(x \leqq u \leqq x')$ is $P(x') - P(x)$. The corresponding area below the horizontal line $y = p(x)$ is $(x' - x)p(x)$. Hence the area R (see Figure 2) is

$$P(x') - P(x) - (x' - x)p(x).$$

Similarly, the area S is $Q(x') - Q(x) - (x' - x)q(x)$.

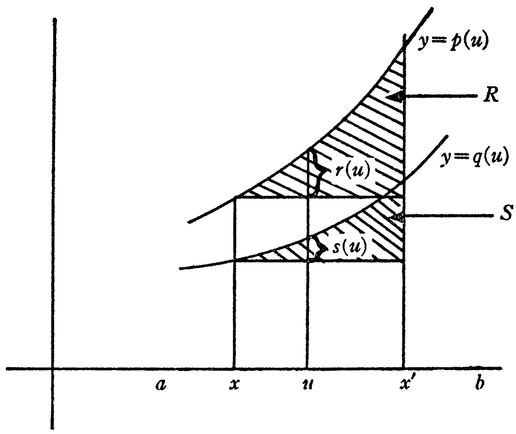


FIG. 2

Since $p-q$ increases on $[x, x']$,

$$p(u) - q(u) > p(x) - q(x) \qquad (x < u)$$

or $r(u)=p(u)-p(x)>q(u)-q(x)=s(u)$. Hence, $R>S$, that is,

$$P(x') - P(x) - (x' - x)p(x) > Q(x') - Q(x) - (x' - x)q(x)$$

and

$$[P(x') - Q(x')] - [P(x) - Q(x)] > (x' - x)[p(x) - q(x)].$$

Thus

$$F(x') - F(x) > (x' - x)f(x).$$

Similarly from $p(u)-q(u)<p(x')-q(x')$, $(u<x')$, we find that

$$F(x') - F(x) < (x' - x)f(x').$$

Since $F(0)=P(0)-Q(0)=0$, $F=P-Q$, is the required area function.

Analytic Proof. To make the previous proof analytic we need to prove two results without an appeal to the figure.

1. For *fixed* x, $P(u)-P(x)-(u-x)p(x)$ is the area function for $p(u)-p(x)$ which vanishes at $u=x$. The analogous statement for Q and q immediately follows.

Proof. For $x\leqq u<u'\leqq x'$, $p(u)(u'-u)<P(u')-P(u)<p(u')(u'-u)$. Add $-p(x)(u'-u)$ throughout. Then

$$\begin{aligned}[p(u) - p(x)](u' - u) &< [P(u') - p(x)u'] - [P(u) - p(x)u] \\ &< [p(u') - p(x)](u' - u).\end{aligned}$$

If we add $-P(x)+xp(x)$ within each of the central brackets, we have the required result.

2. Let $r(u)$, $s(u)$ and $r(u)-s(u)$ be nonnegative and increasing on $x \leqq u \leqq x'$ and $r(x)=s(x)=0$, and let $R(u)$ and $S(u)$ be the area functions of $r(u)$ and $s(u)$ which vanish at $u=x$. Then

$$R = R(x') > S = S(x').$$

Proof. Subdivide $[x, x']$ into $2n$ intervals each of length $\Delta u = (x'-x)/2n$. On each subinterval $[u_{k-1}, u_k]$

$$R(u_k) - R(u_{k-1}) > r(u_{k-1})\Delta u \tag{3}$$

and

$$S(u_k) - S(u_{k-1}) < s(u_k)\Delta u. \tag{4}$$

Change the signs in (4) and add. Then

$$(R - S)(u_k) - (R - S)(u_{k-1}) > [r(u_{k-1}) - s(u_k)]\Delta u.$$

Summing from $k=1$ to $k=2n$,

$$\begin{aligned}(R - S)(x') = R - S &> \left[\sum_1^{2n-1} r(u_k) - \sum_1^{2n} s(u_k)\right]\Delta u \\ &= \sum_1^{2n} [r(u_k) - s(u_k)]\Delta u - r(x')\Delta u \\ &> \sum_n^{2n} [r(u_k) - s(u_k)]\Delta u - r(x')\Delta u \\ &> [r(u_n) - s(u_n)]\, \frac{x' - x}{2} - r(x')\Delta u.\end{aligned}$$

Let

$$r(u_n) - s(u_n)\left[\text{or } r\left(\frac{x + x'}{2}\right) - s\left(\frac{x + x'}{2}\right)\right] = d,$$

and take $\Delta u < (x'-x)d/2r(x')$. Then $R-S>0$ and $R>S$.

5. Volumes. Let a solid be cut by a plane perpendicular to the x-axis at abscissa $x(\geqq 0)$. Let its cross-sectional area $S(x)$ increase with x. If $F(x)$ is the volume between the planes at 0 and x, and $F(x')$ the volume between the planes at 0 and x', we have the double inequality

$$(x' - x)S(x) < F(x') - F(x) < (x' - x)S(x'), \qquad x' > x. \tag{5}$$

For example, if the cross-section is a square of area x^2

$$(x' - x)x^2 < F(x') - F(x) < (x' - x)x'^2. \tag{6}$$

This is the same inequality (1) that arose in connection with the area below the parabola. But in (6), $F(x)$ is the volume of a square pyramid. Hence $F(x) = \frac{1}{3}x^2x = x^3/3$ and $F(x') = x'^3/3$. Therefore, (1) could have been solved by referring to the square pyramid. The area under a parabola is just as elementary as the volume of a pyramid!

It should be clear that problems which involve the distance for a given velocity, the work against a variable force and the like, are equally easy to set up and solve by a simple use of inequalities.

6. Remarks. It is immediately clear that the ordinate function f is the derivative of what we have called the area function F. The method of this paper therefore gives a way of finding derivatives without the use of a limit process.

When f is increasing on $[a, b]$, the graph of $y = F(x)$ is convex and the tangent is characterized by the fact that the curve lies above it except at one point. In fact, for $a < c < x \leqq b$, $(x-c)f(c) < F(x) - F(c)$, that is,

$$F(x) > F(c) + f(c)(x - c).$$

For $a \leqq x < c$, $F(c) - F(x) < (c-x)f(c)$ and $F(x) > F(c) + f(c)(x-c)$. Thus $F(x)$ is above $y = F(c) + f(c)(x-c)$ for all $x \neq c$ on $[a, b]$.

Similarly, if f decreases, $F(x) < F(c) + f(c)(x-c)$ for $x \neq c$.

It is easy to show that if $F(x)$ has a derivative on (a, b) its graph has no finite jumps. It follows that if f itself has a derivative, then f is unique. It is also the case that this method enables one to obtain results which closely parallel those of elementary calculus.

It is frequently emphasized that inequalities should be studied in secondary school, but the applications customarily given do not convince most students of their importance. The fact that results traditionally found by calculus methods can be obtained so easily from the algebra of inequalities immediately opens up significant applications of this algebra. Moreover, it enables the student to handle these applications without the subtleties of limit theory. This is a possible answer to the problem of teaching calculus in secondary school.

It is also our conviction that elementary calculus is now being taught in an oversophisticated way and that we are discouraging many talented students by refined theoretical considerations which are actually unnecessary. If mathematics is as important as we say it is, it is also important that students with modest theoretical appetite should understand it much better than they obviously do. It is therefore incumbent upon us to make mathematics as simple as it can be made without sacrifice of accuracy and power. It is hoped that this paper will contribute to this end.

GEOMETRICAL TRANSFORMATIONS OF VEBLEN-WEDDERBURN SYSTEMS

ROY FEINMAN, Rutgers, The State University

1. Introduction. The purpose of this paper is to generalize some of the familiar geometrical transformations of the euclidean plane to the case of an affine plane coordinatized by a Veblen-Wedderburn system. The approach used is to start with (generalized) line reflections, as in [3], and consider the group G that they generate. Generalized dilatations (Cf. [1]) are also discussed, and a relationship is shown between them and the group G. Finally, it is proved that if the Veblen-Wedderburn system is actually the euclidean plane, the generalized transformations agree with the usual ones.

2. General definitions. A Veblen-Wedderburn system is defined as a linear ternary ring in which the additive loop is an abelian group and the law $(a+b)c = ac+bc$ is satisfied. Details may be found in [4]. The following notation is adopted here: (x, y) denotes a point of the plane, $[n, c]$ denotes the line $y=xn+c$, $[a]$ denotes the line $x=a$. Also, if $ax=b$, we write $x=a\backslash b$; if $xa=b$, we write $x=b/a$. $(a\neq 0)$. The lines $\{[0, a]\}$, $\{[a]\}$, $\{[1, a]\}$ will be referred to as the x, y, and u bundles, respectively. They comprise the coordinatizing three-net of a plane [5].

3. Reflections. The mappings X_c: $(x, y)\to(x, 2c-y)$, Y_c: $(x, y)\to(2c-x, y)$, U_c: $(x, y)\to(y-c, x+c)$ will be called *reflections*.

THEOREM 1. *The reflections are involutory collineations of the x and y bundles and generate a group G. Also, a reflection fixes pointwise a line in the coordinatizing net* (i.e., in the x, y, or u bundle).

The fact that the reflections are involutory mappings with an invariant line, or axis, shows their similarity to euclidean line reflections, which also have these properties. They differ in not being collineations of the entire plane.

Proof. By direct computation it may be shown that the image of a line in the x or y bundle under a reflection is again a line in the x or y bundle. For example, consider the image of a line $[0, a]$ under X_c. We have $[0, a]X_c=\{(x, a)\}X_c=\{(x, 2c-a)\}=[0, 2c-a]$, which is in the x bundle. Further, it is easily shown by computation that the reflections are involutory. Hence they generate the identity mapping and thus a group G. Fixed lines for X_c, Y_c, U_c are $[0, c]$, $[c]$, $[1, c]$ respectively. Thus $[0, c]X_c=[0, 2c-c]=[0, c]$ by the above, etc.

THEOREM 2. *Every element of G has a unique representation $U^iX^jY^kt$, where U, X, Y are the reflections with $c=0$, $i, j, k=1$ or 2, and t is a translation, t: $(x, y)\to(x+a, y+b)$. Further, the additive group of translations T is normal in G with index 2^m, $0\leqq m\leqq 3$.*

Proof. Composing the reflections yields mappings $(x, y)\to(\pm x+a, \pm y+b)$, $(x, y)\to(\pm y+b, \pm x+a)$, which can be expressed uniquely in the given form; for

example, $(x, y) \to (-y + b, x + a)$ may be expressed as $U^1X^2Y^1t$, t: $(x, y) \to (x+b, y+a)$. If we write $G = U^iX^jY^kT$, we see that T is a subgroup of G, given by $i=j=k=2$, and the index of T is 8 or a smaller power of 2. (In $GF(2)$, for example, the index of T in G is unity.) Also, it may be argued that each element of G has a unique representation $tU^iX^jY^k$; the right cosets of T associated with this representation are easily seen to be identical with the corresponding left cosets of T from the other representation, and thus T is normal in G.

THEOREM 3. *In a Veblen-Wedderburn system of additive characteristic* $\neq 2$ *and with commutative multiplication, every translation is the product of at most* 4 *reflections.* (*In the euclidean plane, a translation is the product of* 2 *reflections* [3]).

Proof. We show that t: $(x, y) \to (x+a, y+b) = X_{b/2}X_bY_{a/2}Y_a$. Now $(x, y)X_{b/2}X_b Y_{a/2}Y_a = (x, b/2 + b/2 - y)X_bY_{a/2}Y_a$. But $b/2 + b/2 = [(1 + 1)b]/2 = (2b)/2 = (b\cdot 2)/2 = b$, so that $(x, b/2 + b/2 - y)X_bY_{a/2}Y_a = (x, b- y)X_bY_{a/2}Y_a = (x, y+b)Y_{a/2}Y_a(x+a, y+b) = (x, y)t$.

From this theorem it follows also that in the special case where multiplication is commutative and the additive characteristic $\neq 2$, every element of G is the product of at most 7 reflections; for if $g \in G$, $g = U^iX^jY^kt$, U, X, Y are reflections, and t is the product of at most 4 reflections. The corresponding theorem for the Euclidean plane states that every rigid motion is the product of at most 3 reflections.

4. Dilatations. Now we return to general Veblen-Wedderburn systems. We call the mappings D: $(x, y) \to (xa+b, ya+c)$, D: $(x, y) \to (x/a+b, y/a+c)$, $a \neq 0$, *dilatations*. It is easy to see that the dilatations are one-to-one mappings and that the inverse of a dilatation is a dilatation. For example, the inverse of D: $(x, y) \to (xa+b, ya+c)$ is D^{-1}: $(x, y) \to (x/a - b/a, y/a - c/a)$. Another interesting property, which can be checked by simple computations, is that a translation transformed by a dilatation is again a translation [1].

THEOREM 4. *The dilatations preserve the x, y, and u bundles and the bundle with slope* -1. *Hence they are collineations of the* 4-*net comprised by these bundles.*

Proof. Clearly every dilatation is the product of a dilatation preserving the origin and a translation. By direct computation, it can be verified that the theorem is true for dilatations preserving the origin and for translations; hence it is true in general. For example, a dilatation $(x, y) \to (x/a+b, y/a+d)$ may be decomposed into D_0: $(x, y) \to (x/a, y/a)$ and t: $(x, y) \to (x+b, y+d)$. If, say, D_0 acts on the bundle with slope -1, we have: $[-1, c]D_0 = \{(x, -x+c)\}D_0 = \{(x/a, -(x/a)+c/a)\} = \{(z, -z+c/a)\}$ by repeated use of the distributive law and by substituting $z = x/a$. Hence $[-1, c]D_0 = [-1, c/a]$ for all c and so D_0 preserves the bundle with slope -1. Similar calculations involving t: $(x, y) \to (x+b, y+d)$ complete the proof for this case, and the other cases as well.

THEOREM 5. *A dilatation transformed by an element of G is a dilatation.*

Proof. Let D be a dilatation and g an element of G. Then we may write $gDg^{-1}=U^iX^jY^ktDt^{-1}Y^kX^jU^i$. When the right-hand side operates on (x, y), x and y are not reversed and neither one is negated, since U, X, Y have even degrees in the above expression ($2i$, $2j$, $2k$, respectively). So the only effect of gDg^{-1} on (x, y) is to multiply or divide (on the right) each of x, y by the same element a (the multiplier or divisor in D) and to add some elements r, s to the respective products (quotients). Hence gDg^{-1} is a dilatation.

THEOREM 6. *In a right alternative division ring, G is a collineation group of the entire plane.*

Proof. It suffices to show that U, X, Y and t, t any translation, are collineations of the entire plane. By Theorem 1, these mappings are collineations of the y bundle. Hence consider the lines $\{[m, k]\}$. By using the distributive laws and the right inverse property (Rechtskürzungsregel, [**5**]), it can be verified that U, X, Y are collineations of all these lines. Also, by using only the properties of a Veblen-Wedderburn system, we can compute $[m, k]t=[m, k+b-am]$, t: $(x, y)\rightarrow(x+a, y+b)$. Thus the theorem is proved and we have as a corollary, by the last computation, that T is a collineation group of the plane in any Veblen-Wedderburn system.

It remains to show that our reflections and dilatations become the usual ones in the euclidean plane. It can be checked algebraically that in the euclidean plane, the line reflections with axis in the x, y, or u bundle are in fact the mappings X_c, Y_c, U_c respectively. Also, it is known that the dilatations of the euclidean plane about an arbitrary center are the mappings $z\rightarrow az+d$, z a complex variable, d a complex constant, a real, with reference to the Argand diagram. If we write $z=x+iy$, $d=b+ci$, we get $x\rightarrow ax+b$, $y\rightarrow ay+c$, or $(x, y)\rightarrow(xa+b, ya+c)$. This is sufficient, for in a field, $(x/a+b, y/a+c)=(x\cdot 1/a+b, y\cdot 1/a+c)$.

The author wishes to thank Professor R. Artzy for his advice and encouragement.

Work on this paper was supported by the Undergraduate Science Education Program of the National Science Foundation, Grants GE-97 and GE-1886.

References

1. E. Artin, Geometric Algebra. Interscience Publishers, New York 1957 Ch. 2 Sec. 2.
2. R. Artzy, Linear Geometry, Addison-Wesley, Reading, Mass., 1965.
3. F. Bachmann, Aufbau der Geometrie aus dem Spiegelungsbegriff. Springer Verlag, Berlin, 1959.
4. M. Hall, The Theory of Groups. Macmillan, New York, 1959.
5. G. Pickert, Projektive Ebenen, Springer Verlag, Berlin, 1955.

The Old Professor was asked if there was a particular system of logarithms he preferred. Perhaps he was being redundant but certainly unambiguous in replying with a drawl, "Natural-ly."

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

INTEGRAL SOLUTIONS OF A SYSTEM OF LINEAR EQUATIONS

M. S. CHEEMA, University of Arizona

1. Introduction. There is an important class of problems in which solution vectors of a system of equations or inequalities are required to have integral components. In this note we present a method for solving such a system. In the following X stands for the vector $(x_1, x_2, \cdots, x_n)$.

2. An important theorem (see [2]) gives a sufficient condition for the existence of an integral solution for a set of linear inequalities; in terms of linear forms it is

MINKOWSKI'S THEOREM. *Let*

$$L_i(\mathrm{X}) = \sum_{j=1}^{n} a_{ij}x_j, \qquad 1 \leqq i \leqq n \tag{1}$$

be n linear forms with real coefficients a_{ij} *having* $\Delta = \det [a_{ij}] \neq 0$. *Suppose n positive real numbers* $b_1, \cdots, b_n$ *are such that*

$$\prod_{i=1}^{n} b_i \geqq |\Delta|;$$

then there exists an integral vector C *such that* $L_i(\mathrm{C}) \leqq b_i$, $1 \leqq i \leqq n$.

It may be noticed that the theorem asserts only the existence of a lattice point (integral solution). Furthermore, the condition is by no means necessary. In the case of nonlinear constraints defining a set B the existence of an integral solution is provided if it is possible to find a convex set A centered at the origin with volume $v > 2^n$, contained in B. This is useful for nonlinear problems.

3. Consider a single equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b. \tag{2}$$

It is known that the necessary and sufficient condition for the solvability of (2) in integers is that the greatest common divisor of a_i must divide b (see [1, 4]). In [4] an upper bound is also obtained for the least value, if any, of $\sum_{i=1}^{n} |x_i|$ for the case $b = 1$.

Now we describe a method for solving a system of equations

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad 1 \leqq i \leqq m, \tag{3}$$

where again Δ_i, the greatest common divisor of $a_{i1}, \cdots, a_{in}$ divides b_i for $1 \leqq i \leqq m$. First the system (3) must be checked for consistency.

We take first that one of the equations in which the maximum number of

$a_{ij}=0$, $1\leqq j\leqq n$. Consider then the equation

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where the a_i are so arranged that a_1 is the smallest of all the a_i in absolute value; we can assume $a_1>0$ without loss of generality. Thus

$$x_1 = \frac{b - a_2x_2 - \cdots - a_nx_n}{a_1},$$

and here the right side must be an integer. Reducing all a_i mod a_1 and introducing equivalence mod a_1, we obtain

$$\frac{b - a_2x_2 - \cdots - a_nx_n}{a_1} \equiv \frac{b' - a_2'x_2 - \cdots - a_n'x_n}{a_1} = n_1 \text{ (integer)}.$$

Thus we have to consider the solution of $a_2'x_2+\cdots+a_n'x_n=b'-n_1a_1=b_1$. By rearrangement let a_2' be the smallest of a_i' in absolute value; then

$$x_2 = \frac{b_1 - a_3' - \cdots - a_n'x_n}{a_2'}.$$

Again reducing mod a_2' and introducing equivalence we obtain an equation of the type

$$a_3'x_3 + \cdots + a_n''x_n = -a_2'x_2 + b_1' = b_2,$$

where $a_i\geqq a_i'\geqq a_i''$, $1\leqq j\leqq n$. Therefore the process stops either when all integers $a_i^{(j)}$ reduce mod some $a_k^{(j)}$ to zero except for one, or eventually

$$x_{n-1} = \frac{b_{n-1} - a_n^{(n-2)}x_n}{a_{n-1}^{(n-2)}} \equiv \frac{b_{n-1}' - a_n^{(n-1)}x_n}{a_{n-1}^{(n-2)}} = \text{integer}.$$

This gives x_n in terms of an integral parameter and thus also $x_{n-1},\cdots,x_1$ in terms of $n-1$ parameters because at each stage we introduce an integral parameter. For the system (3) we can substitute these in the remaining equations and now we have to solve $m-1$ equations in $n-1$ unknowns and we can continue the process by again picking up an equation from this system.

We may first proceed to eliminate $m-1$ variables and start with the equation in $n-m+1$ unknowns thus obtained and thereby also obtain the $m-1$ variables also in terms of the integral parameters. Solving inequalities one may introduce slack variables to convert these to equations. The following examples illustrate the method.

Example I. Find integral $\mathbf{X}$ such that

$$\begin{aligned} x_1 - x_3 + 4x_4 &= 3 \\ 2x_1 - x_2 \qquad &= 3 \\ 3x_1 - 2x_2 - x_4 &= 1 \qquad x_2 = -3 + 2x_1 = n_1. \end{aligned}$$

Thus $x_1=(3+n_1)/2\equiv(1+n_1)/2=n_2$.

$$n_1 = 2n_2 - 1$$
$$x_1 = n_2 + 1, \qquad x_2 = 2n_2 - 1$$
$$x_3 = -3n_2 + 14$$
$$x_4 = -n_2 + 4.$$

The general solution is $(n_2+1,\ 2n_2-1,\ -3n_2+14,\ 4-n_2)$ and thus particular solutions are (5, 7, 2, 0), (2, 1, 11, 3), etc. This example is solved by the simplex method in [3] (see page 120).

Example II. Solve in integers:

$$5x + 7y + 8z = 11$$
$$2x - 3y + 6z = 5.$$

From the first equation

$$x = \frac{11 - 7y - 8z}{5} \equiv \frac{1 - 2y - 3z}{5} = l \text{ (integer)}$$
$$1 - 2y - 3z = 5l \quad\text{or}\quad 2y + 3z = -5l + 1$$
$$y = \frac{-5l + 1 - 3z}{2} \equiv \frac{-l + 1 - z}{2} = k \text{ (integer)}.$$

Thus, $z=-l-2k+1$, $y=-l+3k-1$, $x=2+3l-k$. Substituting in the second equation and applying the same technique one obtains the solution

$$x = 16 + 66n, \qquad y = -3 - 14n, \qquad z = -6 - 29n,$$

where n may be given any integral value.

4. Using transference theorems [2, pp. 75–76], we can make information about the primal problem give information about the dual problem. Let $\|\ \|$ stand for the distance from the nearest integer. Let

$$L_j(\mathrm{X}) = \sum_{i=1}^{m} \theta_{ij}x_i, \qquad 1 \leqq j \leqq n \tag{4}$$

be n linear forms in m variables and let

$$M_i(\mathrm{U}) = \sum_{j=1}^{n} \theta_{ji}u_j \tag{5}$$

be the transposed set of m linear forms in n variables.

Using Minkowski's theorem, we can show (see [2]) that:

THEOREM. *There is always an integral* $\mathrm{X}\neq 0$ *such that*

$$\|L_j(\mathrm{X})\| \leqq C, \qquad (1 \leqq j \leqq n,\ |x_i| \leqq X,\ 1 \leqq i \leqq m) \tag{6}$$

for any $x>1$ *and* $c=x^{-m/n}$.

It can also be shown that if (6) is solvable with $X \neq 0$ for some X and some C much smaller than $X^{-m/n}$, then the transposed set,

$$\|M_i(\mathbf{U})\| \leqq D, \qquad |U_j| \leqq U \tag{7}$$

is solvable with $\mathbf{U} \neq 0$ for some U depending on X and C and some D much smaller than the expected $U^{-n/m}$. There is also a relation between the homogeneous problem and the corresponding inhomogeneous problem of solving

$$\|L_j(\mathbf{X}) - \alpha_j\| \leqq C_1, \qquad |x_i| \leqq X_1 \tag{8}$$

with integral $\mathbf{X}$ for given $\boldsymbol{\alpha}$.

The author wishes to thank the referee for his suggestions. This work was supported by National Science Foundation Grant GP24343.

References

1. E. Artin, Theory of Algebraic Numbers, Göttingen Lectures, 1956.
2. J. W. S. Cassels, An Introduction to Diophantine Approximations, Cambridge University Press, 1957.
3. D. Gale, The Theory of Linear Economic Models, McGraw-Hill, New York, 1960.
4. S. S. Pillai, On a Linear Diophantine Equation, Proc. Indian Acad. Sciences, Sect. A, 12 (1940) 199–201.

ON THE COMPLETENESS OF THE TOTITIVES OF A NATURAL NUMBER

V. O. S. Olunloyo and A. D. Weiss, University of Ife, Nigeria

1. Introduction. Following Sylvester [1], the natural numbers $<m$ and relatively prime to m are called the totitives of m. The others $<m$ and not relatively prime to m are called non-totitives. It is natural to inquire as to which natural numbers m have the property that the totitives form a complete residue system mod $\phi(m)$. Such numbers shall be referred to as admissible. That every prime p is admissible is trivial since $\phi(p)=p-1$, the totitives being $1, 2, \cdots, p-1$. The case $m=1$ is also, by convention, trivial.

In this paper, we prove, employing no more than the Fundamental Theorem of Arithmetic, that 15 is the only composite admissible number. This property of 15 may thus be regarded as one furnishing an alternative characterisation of the primes.

To be definite, any totitive t of m may be written in the form $t=q\phi(m)+r$, q and r being natural numbers such that $0 \leqq r \leqq \phi(m)-1$. t will then be said to represent residue class r. For m to be admissible, every class must be represented exactly once. From now on we assume that $m>2$.

2. Theorem 1. *Any $m>2$ is admissible only if $(m, \phi(m))=1$.*

Proof. If $(m, \phi(m)) \neq 1$, then $(m, h\phi(m)) \neq 1$ for any integer h. This implies Class 0 will not be represented.

COROLLARY 1.1. *$m > 2$ is admissible only if m is odd. $\phi(m)$ is even for $m > 2$ and so $(m, \phi(m))$ will be $\geqq 2$, if m is even.*

COROLLARY 1.2. *A composite m is admissible, only if it is a product of distinct primes.*

For, if the canonical decomposition of m is

$$m = \prod_{i=1}^{k} p_i^{\alpha_i}, \qquad \alpha_i \geqq 1, \qquad p_i \geqq 3 \quad \text{and} \quad p_i < p_{i+1},$$

then

$$\phi(m) = \prod_{i=1}^{k} p_i^{\alpha_i - 1}(p_i - 1).$$

Thus if for some i, $\alpha_i \geqq 2$, $p_i | \phi(m)$ and $(m, \phi(m)) \geqq p_i \neq 1$. From now on we assume that

$$m = \prod_{i=1}^{k} p_i, \qquad 3 \leqq p_i < p_{i+1}. \tag{2.1}$$

3. We find it convenient here to define an arithmetic function, denoted by N_m, which we use in the course of the paper.

$$N_m = 2\phi(m) - m.$$

THEOREM 2. *m is admissible, only if $N_m > 0$, i.e.,*

$$\phi(m) > m/2. \tag{3.1}$$

Proof. If $2\phi(m) < m$, then by Theorem 1 and Corollary 1.1, $(m, 2\phi(m)) = (m, 2) = 1$. Thus $2\phi(m)$ and $\phi(m)$ would both be totitives congruent mod $\phi(m)$ contrary to the necessity of class 0 being represented at most once. Also an admissible m cannot equal $2\phi(m)$ which is even. Hence Theorem 2, which incidentally implies the existence of at least one x such that $m - \phi(m) \leqq x \leqq \phi(m)$.

4. THEOREM 3. *m is admissible only if every x such that $m - \phi(m) \leqq x \leqq \phi(m)$ is a totitive.*

Proof. We make a few preliminary general observations on the general distribution of the totitives and non-totitives of any natural number. If t is a totitive of m, so is $m - t$ because in general $(m, t) = (m, m - t)$, and if $1 \leqq t < m$ then $1 \leqq m - t < m$. Thus the totitives can clearly be partitioned in pairs according as to whether they sum to m. Also, if n is a non-totitive of m then so is $m - n$. For any admissible m, it appears that one of the two situations (i) and (ii) below presents itself.

(i) There exists no x such that $m-\phi(m)<x<\phi(m)$. In this case $N_m=1$ and $m-\phi(m)$ and $\phi(m)$ are consecutive totitives, and there is nothing to prove.

(ii) There exists at least one x such that $m-\phi(m)<x<\phi(m)$ in which case $N_m>1$. In this case, $x+\phi(m)>m$ and $x-\phi(m)<0$ $\therefore x\not\equiv y \pmod{\phi(m)}$, $1\leqq y\leqq m$ unless $x=y$. Since $x<\phi(m)$ the necessity of representing class x demands that x be a totitive. Theorem 3 follows.

COROLLARY 3.1. *m is admissible only if $N_m<p_1-1$.*

If $N_m\geqq p_1-1$, the consecutive numbers from $m-\phi(m)$ to $\phi(m)$ inclusive will constitute at least p_1 consecutive numbers and at least one of these, a multiple of p_1 in fact, would be non-totitive contrary to Theorem 3. The inequality would now be replaced by an equation.

5. THEOREM 4. *m is admissible only if $\phi(m)+1=\frac{1}{2}(m+p_1)$.*

Proof. Since 1 is in general always a totitive, $\phi(m)+1$ must be a non-totitive if m is admissible. $m-\phi(m)-1$ must be a non-totitive too. We show first that $\phi(m)+1\leqq\frac{1}{2}(m+p_1)$. Suppose that m is admissible and $\phi(m)+1>\frac{1}{2}(m+p_1)$. Now $\frac{1}{2}(m+p_1)>\frac{1}{2}m$ and $m-\phi(m)<m/2$ since by Theorem 2, $\phi(m)>\frac{1}{2}m$. Hence $m-\phi(m)\leqq\frac{1}{2}(m+p_1)\leqq\phi(m)$. Since $p_1|\frac{1}{2}(m+p_1)$ we have a non-totitive x such that $m-\phi(m)\leqq x\leqq\phi(m)$, contrary to Theorem 3. Suppose again that m is admissible and $\phi(m)+1<\frac{1}{2}(m+p_1)$. Then $m-\phi(m)-1>m-\frac{1}{2}(m+p_1)$, i.e., $m-\phi(m)-1>\frac{1}{2}(m-p_1)$. Since $\phi(m)+1>m/2$, and therefore $m-\phi(m)-1<m/2$; on combining the inequalities we have

$$\frac{1}{2}(m-p_1)<m-\phi(m)-1<\phi(m)+1<\frac{1}{2}(m+p_1). \tag{5.1}$$

But $m-\phi(m)-1$ and $\phi(m)+1$ are both non-totitive with at least one same common factor, say p_j, in common with m. Their difference must therefore be $\geqq p_j$. From (5.1), however, their difference must be less than $\frac{1}{2}(m+p_1)-\frac{1}{2}(m-p_1)=p_1$. Since every $p_j\geqq p_1$ we arrive at a contradiction, and the proof is complete.

COROLLARY 4.1. *m is admissible only if $p_1|\phi(m)+1$.*

Since $\phi(m)+1=\frac{1}{2}(m+p_1)$ and $p_1|\frac{1}{2}(m+p_1)$, this follows immediately.

COROLLARY 4.2. *m is admissible only if $N_m=p_1-2$.*

Since $\phi(m)+1=\frac{1}{2}(m+p_1)$, $2\phi(m)+2=m+p_1$ $\therefore N_m=p_1-2$.

6. THEOREM 5. *A composite number is admissible only if it is a multiple of* 3.

Proof. If $m=p_1p_2\cdots p_k$, $p_1\geqq 3$ then every number between 1 and p_1-1 inclusive is a totitive and thus, in particular, p_1-2 is a totitive. We intend to show that $\phi(m)+p_1-2$ must be a non-totitive. It suffices to show that $\phi(m)+p_1-2<m$. Now

$$m-p_k=p_k(p_1\cdots p_{k-1}-1)>p_k(p_1-1)(p_2-1)\cdots(p_{k-1}-1)>\phi(m);$$

hence $\phi(m)+p_k<m$ and, *a fortiori*, $\phi(m)+p_1-2<m$. Moreover,

$$\phi(m)+p_1-2=\phi(m)+1+(p_1-3)=\tfrac{1}{2}(m+p_1)+(p_1-3)=\tfrac{1}{2}(m+3p_1-6)$$

by Theorem 4. Thus there exists a p_j, $1\leqq j\leqq k$ such that $p_j\,|\,\tfrac{1}{2}(m+3p-6)$; hence $p_j\,|\,3(p_1-2)$ and, since p_j cannot divide p_1-2, $p_j\,|\,3$, whence $p_j=3$ and, incidentally, $j=1$. Theorem 5 follows.

COROLLARY 5.1. *m is admissible only if $N_m=1$.*

This follows immediately from Corollary 4.2.

7. THEOREM 6. *A composite number is admissible only if it is a multiple of* 15.

Proof. For m odd, 2 is always a totitive and therefore, if a composite m is admissible then $\phi(m)+2$ must be a non-totitive. Then

$$\phi(m)+2=\tfrac{1}{2}(m+p_1)+1=\tfrac{1}{2}(m+p_1+2).$$

There exists a p_j therefore, with $1\leqq j\leqq k$, such that $p_j\,|\,m+p_1+2$ and, since $p_j\,|\,m$, $p_j\,|\,p_1+2$. Since p_1 cannot divide p_1+2, $p_j>p_1$ and $p_j\geqq p_1+2$. Since p_j cannot be greater than p_1+2 and divide it, $p_j=p_1+2$. Thus, by Theorem 5, $p_j=3+2=5$ must be a factor of m. Combining this result with Theorem 5 yields Theorem 6.

8. We now characterise admissibility.

THEOREM 7. *A composite number is admissible if and only if $N_m=1$ and there exists no pair of non-totitives congruent* mod $\phi(m)$.

Proof. To prove necessity, in view of Corollary 5.1, it remains to show that it is necessary that there exists no pair of non-totitives congruent mod $\phi(m)$. If m is admissible, suppose that there exists at least one pair of congruent non-totitives a, b, where $b=a+\phi(m)$. But since $m<2\phi(m)$, class a will thus not be represented contrary to hypothesis that m is admissible. To prove sufficiency, given $N_m=1$ and that there exists no pair of congruent non-totitives, suppose that m is not admissible. Then at least one class, say class α, will not be represented. If $1\leqq\alpha\leqq m-\phi(m)-1$ then $(\alpha, m)\neq 1$ and also in particular $(\alpha+\phi(m), m)\neq 1$. But this gives rise to a pair of congruent non-totitives contrary to hypothesis. If, moreover, $m-\phi(m)\leqq\alpha\leqq\phi(m)$ and $N_m=1$, then $\alpha=m-\phi(m)$ or $\alpha=\phi(m)$. If $\alpha=\phi(m)$ then $(m, \phi(m))\neq 1$, $(m, m-\phi(m))\neq 1$, and there exists $p_j\geqq 3$ such that $p_j\,|\,N_m$ contrary to $N_m=1$. The same holds if $\alpha=m-\phi(m)$, and the proof is thus complete.

THEOREM 8. *There is no composite admissible number greater than* 21.

Proof. If m is admissible then $N_m=1$ i.e. $\phi(m)=\tfrac{1}{2}(m+1)$. Then $\phi(m)+10=\tfrac{1}{2}(m+21)$, which is less than m if m is greater than 21. If an admissible m is greater than 21, then both 10 and $\phi(m)+10$ would be less than m. They are congruent mod $\phi(m)$. Moreover, since $3\,|\,m$ and $3\,|\,\phi(m)+1$ it follows that $3\,|\,\phi(m)+10$ and so $\phi(m)+10$ would be a non-totitive. Again since $5\,|\,m$ and $5\,|\,10$, 10

would be a non-totitive. Thus for a composite admissible $m > 21$, there always exists two congruent non-totitives which violates Theorem 7. This contradiction shows that there is no admissible composite number greater than 21. We thus come to the main result.

THEOREM 9. *The only composite admissible natural number is* 15.

Proof. If a composite $m > 2$ is admissible then $15 | m$ and $m \leqq 21$. Thus clearly m can only be 15. It only remains to examine 15 for the property. $\phi(15) = \phi(3)\,\phi(5) = 8$, the totitives being 1, 2, 4, 7, 8, 11, 13, 14. This is a complete residue system mod 8. Indeed the least nonnegative residues are 1, 2, 4, 7, 0, 3, 5, 6 respectively. This establishes the theorem.

9. Since the only admissible numbers, apart from unity, are 15 and the primes, the notion of admissibility appears to be logically almost equivalent to that of primality. The idea of admissibility involves basically, however, only the ideas of relative primality and complete system of residues, both of which derive directly from Euclid's algorithm. Certainly it need not involve the notion of primality directly. Excluding the unique admissible number which has admissible factors, we are left with the primes. Thus it becomes clear that with slight modifications the notion of a prime may be replaced with the idea of admissibility. Although one cannot necessarily foresee any advantage in this line of action, it stands to be considered a potential alternative idea with which one could build a relatively consistent theory of numbers. The bearing of the theorem to the theory of the exponent to which a number belongs modulo m needs consideration. Also since the totitives of m form an Abelian group under multiplication mod m, there is some need to find out whatever may be peculiar about the group of the totitives of 15. These logical and algebraic implications, together with the bearing on other problems in number theory, would form the subject of a subsequent paper.

Reference

1. L. E. Dickson, History of the Theory of Numbers, Vol. 1, Chelsea, New York 1952, p. 124.

THE DECOMPOSITION OF A RATIONAL PRIME IDEAL IN CYCLOTOMIC FIELDS

G. BACHMAN, Polytechnic Institute of Brooklyn

In this note, we wish to give an alternate proof of the way a rational prime ideal decomposes in the ring of integers of a cyclotomic field (see Theorem 2 below). This can be achieved along the lines developed, e.g. in [3] chapter 8, by using the ramification theory and determining first that it is those and only those primes dividing the field discriminant that ramify. The theorem can also be obtained in a more abstract manner, not making full use of the ramification theory, but just determining a finite set of primes from which the ramified ones must come, and by using the Frobenius automorphism (see, e.g. [5] chapter 7). We wish to obtain the theorem in the valuation spirit of [1] and by using essentially local arguments, avoiding completely the ramification theory.

In order to achieve our goal, we shall make use of the following theorem (see e.g. [5] p. 58, [1] p. 118, and [2] p. 133).

THEOREM 1. *If $K=k(\alpha)$ is a simple separable algebraic extension of k and if $|\,|$ is a rank one valuation of k, then there are as many distinct extensions of $|\,|$ to K as there are distinct irreducible factors of $f(x)$ over $\hat{k}$, where $f(x)$ is the irreducible polynomial satisfied by α over k and where $\hat{k}$ is the completion of k. Moreover, the global degree of the extension is equal to the sum of the local degrees.*

Thus, using the valuation approach to ideal theory as presented, e.g. in [1] chapter 8, in order to see how (p) factors in the ring of integers of the cyclotomic field $Q(\zeta)$, we must consider the factorization of the cyclotomic polynomial $\Phi_m(x)$ satisfied by ζ over the p-adic field Q_p.

In order to achieve this we recall now some facts from algebra concerning the factorization of $\Phi_m(x)$ over GF(p), where we assume first that $(m, p)=1$. Consider first the polynomial x^m-1 over GF(p). In order to determine the splitting field of this polynomial over GF(p), we must determine the smallest extension field E in which x^m-1 has m distinct roots. These roots in E form a group, but $E^*=E-\{0\}$ is a cyclic group. Hence we must find the smallest extension field E such that E^* has a cyclic subgroup of order m. For E^* to have a cyclic subgroup of order m, we must just have $m\mid \operatorname{ord} E^*$. Thus, we must find the smallest n such that $m\mid p^n-1$, and then GF(p^n) is the desired splitting field. We denote this n by $e_p(m)$. Next, consider $\Phi_m(x)$ over GF(p). If we form the splitting field of $\Phi_m(x)$ over GF(p), then this field contains the primitive mth roots of unity and hence all the mth roots of unity. If we simply adjoin to GF(p) one root of $\Phi_m(x)$, then we clearly obtain the splitting field of x^m-1. But, GF$(p^{e_p(m)})$ is the splitting field of x^m-1. If $g(x)$ is any irreducible factor of $\Phi_m(x)$ over GF(p), then adjoining a root of $g(x)$ also yields the splitting field of x^m-1; i.e. GF$(p^{e_p(m)})$. Hence, $\deg g(x)=e_p(m)$, and, since this is true for any irreducible factor of $\Phi_m(x)$, we get the following factorization of $\Phi_m(x)$ over GF(p):

$$(*)\qquad \Phi_m(x) = \prod_{i=1}^{l} \psi_{d_i}(x),$$

where $l=\varphi(m)/e_p(m)$ and where the degree of each $\psi_{d_i}(x)$ is $e_p(m)$.

In case m and p are not relatively prime, we write $m=p^\alpha r$, where $(r, p)=1$ and then use the fact that

$$\Phi_{p^\alpha r}(x) = \Phi_r(x)^{\varphi(p^\alpha)}$$

which follows readily from the relation $\Phi_m(x)=\prod_{d\mid m}(x^d-1)^{\mu(m/d)}$, given in [4] p. 114.

Now let us return to the problem at hand, namely, the factorization of $\Phi_m(x)$ over Q_p. Assume first that $(m, p)=1$. Over GF(p), $\Phi_m(x)$ factors according to (*), where the $\psi_{d_i}(x)$ are distinct irreducible factors. Thus by Hensel's lemma

(see e.g. [5] p. 45 or [4] p. 248), $\Phi_m(x)$ factors into l irreducible factors over Q_p. But the global degree, $\varphi(m)$, is the sum of the local degrees, i.e. $\varphi(m) = \sum_{i=1}^{l} n_i$, and each $n_i = e_i f_i$, where e_i is the ramification index, and f_i the residue class degree of the extension. But $f_i = \deg \psi_{d_i}(x) = e_p(m)$. Thus,

$$\varphi(m) = e_p(m) \sum_{i=1}^{l} e_i,$$

and if any $e_i > 1$, then $\varphi(m) > e_p(m)l$, contradicting (*). Thus each $e_i = 1$, and

$$(p) = y_1 y_2 \cdots y_l,$$

where the y_i are prime ideals in the ring of integers of $Q(\zeta)$, and where $l = \varphi(m)/e_p(m)$.

Finally suppose that $(m, p) \neq 1$. Write $m = p^\alpha m'$ where $(p, m') = 1$. Then, over GF(p),

$$\Phi_m(x) = \prod_{i=1}^{l'} \psi_{d_i}(x)^{\varphi(p^\alpha)}$$

where the degree of each $\psi_{d_i}(x) = e_p(m')$ and $l' = \varphi(m')/e_p(m')$. Thus the residue class degree $f_i = e_p(m')$; whereas the local degree n_i is, by Hensel's lemma, certainly less than or equal to $e_p(m')\varphi(p^\alpha)$, so $e_i \leqq \varphi(p^\alpha)$. If we can show that there exist precisely l' extensions of the p-adic valuation, then

$$\varphi(m) = \sum_{i=1}^{l'} e_i f_i = e_p(m') \sum_{i=1}^{l'} e_i$$

and if any $e_i < \varphi(p^\alpha)$, then we get

$$\varphi(m) < e_p(m') l' \varphi(p^\alpha) = \varphi(m')\varphi(p^\alpha),$$

a contradiction. Hence each $e_i = \varphi(p^\alpha)$, and it remains to show only that there are precisely l' extensions.

Consider then $\psi_{d_i}(x)^{\varphi(p^\alpha)}$ and let the corresponding polynomial given by Hensel's lemma over the valuation ring V of Q_p be $h_i(x)$. If

$$h_i(x) = f_1(x)^{k_1} \cdots f_t(x)^{k_t}$$

where the $f_i(x)$ are distinct irreducible polynomials over V, then the image $\bar{f}_i(x)$ must be a power of an irreducible polynomial in GF(p)[x]; whence, by the unique factorization theorem in GF(p)[x], we must have $\bar{f}_i(x) = \psi_{d_i}(x)^{s_i}$. If $i \neq j$, however, then there exist polynomials $g_i(x)$, $g_j(x)$ in $V[x]$ such that

$$1 = g_i(x) f_i(x) + g_j(x) f_j(x),$$

so

$$1 = \bar{g}_i(x)\bar{f}_i(x) + \bar{g}_j(x)\bar{f}_j(x),$$

or

$$1 = \bar{g}_i(x)\psi_{d_i}(x)^{s_i} + \bar{g}_j(x)\psi_{d_i}(x)^{s_j}.$$

Therefore, we must have $t=1$, and $h_i(x)=f_1(x)^{k_1}$, which shows that there are precisely l' extensions and completes the proof. We have established the following result.

THEOREM 2. (a) *If* $(p, m)=1$, *then in the cyclotomic field* $Q(\zeta)$ *of degree* $\varphi(m)$, $(p)=y_1y_2 \cdots y_l$, *where the* y_i *are prime ideals of* $Q(\zeta)$ *and* $l=\varphi(m)/e_p(m)$, *where* $e_p(m)$ *is the smallest* n *such that* $p^n \equiv 1 \pmod{m}$.

(b) *If* $(p, m) \neq 1$, *let* $m=p^\alpha m'$, $(m', p)=1$, *then* $(p)=(y_1y_2 \cdots y_{l'})^{\varphi(p^\alpha)}$, *where the* y_i *are primes and* $l'=\varphi(m')/e_p(m')$.

References

1. E. Artin, Theory of Algebraic Numbers, Lecture Notes, Göttingen, 1956.
2. G. Bachman, Introduction to p-adic Numbers and Valuation Theory, Academic Press, New York, 1964.
3. H. B. Mann, Introduction to Algebraic Number Theory, Ohio State University Press, 1964.
4. B. L. van der Waerden, Modern Algebra Vol. I, Ungar New York, 1949.
5. E. Weiss, Algebraic Number Theory, McGraw-Hill, New York, 1963.

ON TRAVERSING GRAPHS

D. J. TROY, Ohio State University

Let $G=(V, U)$ be a finite graph, V its set of vertices, U its set of edges. For any x in V denote by $d(x)$ the degree or valency of x, that is, the number of edges incident to x, loops counted twice. Denote by $V(G, n)$ the number of vertices of degree n. Define a path of length k to be a $(2k+1)$-tuple of the form $(x_0, \epsilon_1, x_1, \epsilon_2, \cdots, x_{k-1}, \epsilon_k, x_k)$, where, for $i=0, 1, 2, \cdots, k$, x_i is in V and, for $i=1, 2, \cdots, k$, ϵ_i is an edge incident to both x_{i-1} and x_i. A path is a cycle if $x_k=x_0$. Further, define a cycle to be proper if $\epsilon_i \neq \epsilon_{i+1}$ for $i=1, 2, \cdots, k-1$ and $\epsilon_k \neq \epsilon_0$.

In any finite connected graph there exists a cycle in which every edge appears exactly twice, once in each direction [1, p. 41]. Obviously such a cycle can not be proper if $V(G, 1)>0$. It is not difficult to find a proper cycle with this property if G is a connected Euler graph having at least one vertex x such that $d(x)>2$. Define a covering of a graph G to be a proper cycle in which every edge appears exactly twice, once in each direction. A class of graphs for which no covering exists is given by the following result.

THEOREM. *Let* $V(G, n)=0$ *for all* $n \geqq 4$ *and* $V(G, 3) \equiv 0 \pmod{4}$. *Then no covering of* G *exists.*

Proof. Let us first note the following: (A) Let $d(x)=3$ and ϵ_i $(i=1, 2, 3)$ be the edges incident to x. If a proper cycle P includes each of these edges exactly

twice, once in each direction, then these edges are distinct; that is, there is no loop at x. Further, if y_i $(i=1, 2, 3)$ is the vertex connected to x by ϵ_i, then not both $(y_1, \epsilon_1, x, \epsilon_2, y_2)$ and $(y_2, \epsilon_2, x, \epsilon_1, y_1)$ appear as subpaths of P. (B) Let $d(x) = d(y) = d(w) = d(z) = 3$, x, y, z and w four distinct vertices. Let x and y be connected by two edges α and μ. Let λ connect x and w and β connect y to z. Then $G=(V, U)$ has a covering if and only if $G'=(V-\{x, y\}, U-\{\alpha, \mu, \lambda, \beta\})$ has a covering.

(A) should cause the reader little difficulty. To show that (B) holds, suppose first that G has a covering P. In order that both α and μ appear exactly twice, once in each direction, it is necessary that the paths $X=(w, \lambda, x, \alpha, y, \mu, x, \lambda, w)$ and $X'=(z, \beta, y, \alpha, x, \mu, y, \beta, z)$ or X and X' with α and μ interchanged appear as subpaths of P. By (A), the edge preceding X and the edge following X are distinct. Hence X and, by the same argument, X' may be removed from P leaving a covering of G'. Into a covering of G', X and X' may be inserted to form a covering of G.

We assume that G is connected and that $V(G, 1)=0$, otherwise the result is trivial. Let $V(G, 3)=4k$. If $k=0$, all vertices are of degree 2 and G is a simple closed path and not coverable.

Let $m>0$ and assume that G has no covering if $k<m$. Now consider the case in which $k=m$ and without loss of generality assume that $V(G, 2)=0$.

If G has a separating edge σ, then the graph $G^*=(V, U-\{\sigma\})$ has two disjoint components G' and G'', where $V(G', 2)=V(G'', 2)=1$ and $V(G', 3)+V(G'', 3)=4m-2$. A covering of G if considered to start with the edge σ must cover one of G' or G'' before the second appearance of σ and then cover the other. In G', however, the vertices of degree 3 are the only vertices of odd degree. It is known that in any finite graph the number of vertices of odd degree is even [1, p. 10]; hence $V(G', 3)$ is even. The same argument applies to G''. Therefore, one of $V(G', 3)$ or $V(G'', 3)\equiv 0 \pmod 4$ and is less than $4m$. It follows that not both G' and G'' are coverable; hence G is not coverable.

Since $d(x)=3$ for all x and $V(G, 3)=4m>2$, any given pair of vertices can be connected by at most two edges. Let x and y be connected by two edges, and w and z be respectively the other vertex connected to x and the other vertex connected to y; $w\neq z$, for otherwise there is a separating edge. Using (B) and its terminology, G has a covering if and only if G' has a covering. But $V(G', 2)=2$ and $V(G', 3)=4m-4$; hence G' has no covering.

It may now be assumed that G has no separating edges or multiple edges. To arrive at a final contradiction assume that P is a covering of G. Let h be the smallest integer such that there exists an elementary cycle of length h within P; $h\geqq 3$. Let $(x_1, \epsilon_1, x_2, \epsilon_2, x_3, \epsilon_3, \cdots, \epsilon_h, x_1)$ be such a cycle. Let ϵ_0 be the third edge incident to x_1 and x_0 be the other vertex incident to ϵ_0. For $i=1, 2, \cdots, h$, $x_0\neq x_i$. Consider P as starting with the directed edge (x_0, ϵ_0, x_1) and ending with a directed edge (q, β, x_0); let π be the third edge incident to x_0 and p the other vertex on π; let r be the third vertex connected to x_2 by an edge λ; let s be the

third vertex connected to x_3 by an edge μ. P is then of the form

$$(x_0, \epsilon_0, x_1, \epsilon_1, x_2, \epsilon_2, x_3, \epsilon_3, \cdots, \epsilon_h, x_1, \epsilon_0, x_0, \pi, p, \cdots, \\ s, \mu, x_3, \epsilon_2, x_2, \lambda, r, \cdots, q, \beta, x_0).$$

Let $M=(x_3, \epsilon_3, \cdots, \epsilon_h, x_1)$, $N=(x_0, \pi, p, \cdots, s, \mu, x_3)$ and $Q=(x_2, \lambda, r, \cdots, q, \beta, x_0)$; so that

$$P = (x_0, \epsilon_0, x_1, \epsilon_1, x_2, \epsilon_2, x_3) + M + (x_1, \epsilon_0, x_0) + N + (x_3, \epsilon_2, x_2) + Q.$$

Let $P'=(x_1, \epsilon_1, x_2)+Q+N+M$. P' is a proper cycle that includes neither of the edges ϵ_0 or ϵ_2 but does include every other edge of G exactly twice, once in each direction. Hence the graph $(V, U-\{\epsilon_0, \epsilon_2\})$, which has $4m-4$ vertices of degree 3, is coverable, and this is the contradiction sought.

The author has found many examples of homogeneous graphs of degree 3 with $4k+2$ vertices which are coverable and examples of some that are not. None of his noncoverable examples are triconnected graphs. Biconnectedness is not sufficient but as can be seen from remark (B) triconnectedness is not necessary.

Reference

1. Oystein Ore, Theory of Graphs, Amer. Math. Soc. Colloq. Publ., 38 (1962).

SIMILAR INVOLUTORY MATRICES (mod p^m)

JOEL BRAWLEY, JR., North Carolina State University

If A is an $n\times n$ matrix over the field GF(p^s), where p is an odd prime, it is known that A is involutory $(A^2=I)$ if and only if A is similar to a canonical matrix $J_t=\text{diag}\,(I_t, -I_{n-t})$, $0\leqq t\leqq n$, for some unique integer t [1; p. 518]. In this paper we prove a corresponding result can be obtained for an involutory matrix over the integers taken modulo p^m. This is the essence of Theorem 1. Using this result we obtain the number of involutory matrices modulo p^m by the method of Hodges [1]. The number is seen to agree with that obtained by Reiner [3], and with a similar result of Levine and Korfhage [2].

All matrices considered below will be understood to have integral elements. When $A=(a_{ij})$ and $B=(b_{ij})$ are such matrices, by $A\equiv B \pmod q$ we mean that $a_{ij}\equiv b_{ij} \pmod q$ for all i, j; i.e., $A=B+qC$ for some integral matrix C. We say that A is involutory modulo q if $A^2\equiv I \pmod q$, and by the statement "the matrix B modulo q," we mean the elements b_{ij} of B are taken so that $0\leqq b_{ij}<q$.

THEOREM 1. *If A is an $n\times n$ matrix with integral elements, and p is an odd prime, then A is involutory modulo p^m if and only if there exist matrices Q, P and a unique integer t $(0\leqq t\leqq n)$ such that*

(1) $\quad QP\equiv I \pmod{p^m}$, $\qquad$ (2) $\quad QAP\equiv J_t \pmod{p^m}$,

where $J_t=\text{diag}\,(I_t, -I_{n-t})$.

Proof. The sufficiency of the theorem is immediate for if (1) and (2) hold, then from (1) we have $|Q|\ |P| \equiv 1 \pmod{p^m}$ so that P has a (two sided) inverse modulo p^m; namely $P^{-1} = |Q|$ adj P. Multiplying both sides of (1) on the right by P^{-1} we find that $Q \equiv P^{-1} \pmod{p^m}$; hence Q is the inverse of P modulo p^m. Thus, $QP \equiv PQ \equiv I \pmod{p^m}$ so that from (2) we have $A \equiv PJ_tQ \pmod{p^m}$ and $A^2 \equiv (PJ_tQ)(PJ_tQ) \equiv PJ_t^2Q \equiv PIQ \equiv I \pmod{p^m}$.

Conversely, let us assume that A is involutory modulo p^m. Then $A^2 - I = p^mC$ for some C; hence for $i = 1, 2, \cdots, m$ we have that $A^2 \equiv I \pmod{p^i}$. In particular for $i = 1$, this implies the existence of matrices P_1 and Q_1 such that $Q_1P_1 \equiv I \pmod{p}$ and $Q_1AP_1 \equiv J_t \pmod{p}$ for some (unique) t. This is because the congruence relation taken modulo p yields a field. We make the induction assumption of the existence of matrices P_{k-1}, $Q_{k-1} (2 \leqq k \leqq m)$ such that

$$Q_{k-1}P_{k-1} = I + p^{k-1}M \equiv I \pmod{p^{k-1}}, \tag{3}$$

$$Q_{k-1}AP_{k-1} = J_t + p^{k-1}N \equiv J_t \pmod{p^{k-1}}. \tag{4}$$

Then we let $Q_k = Q_{k-1} + p^{k-1}X$ and $P_k = P_{k-1} + p^{k-1}Y$, with X arbitrary modulo p and Y defined by $Y \equiv -P_{k-1}M - P_{k-1}XP_{k-1} \pmod{p}$. By direct multiplication

$$Q_kP_k \equiv I \pmod{p^k} \tag{5}$$

and, using (3) and (4), we see that

$$Q_kAP_k \equiv J_t + p^{k-1}[N - J_tM - J_tXP_{k-1} + XAP_{k-1}] \pmod{p^k}. \tag{6}$$

Since $A^2 \equiv I \pmod{p^k}$ and $Q_kP_k \equiv P_kQ_k \equiv I \pmod{p^k}$, we have the identity $(Q_kAP_k)^2 \equiv I \pmod{p^k}$. This identity is true for arbitrary X in (6), and hence, by noting it for $X = 0$, we find

$$[J_t + p^{k-1}(N - J_tM)]^2 = I \pmod{p^k},$$

or equivalently,

$$J_tN - M \equiv J_tMJ_t - NJ_t \pmod{p}. \tag{7}$$

Thus, taking $X \equiv (r+1)[J_tN - M]Q_{k-1} \pmod{p}$, where $p = 2r+1$, we use (7) in (6) to obtain

$$\begin{aligned} Q_kAP_k &\equiv J_t + p^{k-1}[N - J_tM - (r+1)J_t(J_tN - M)Q_{k-1}P_{k-1} \\ &\qquad + (r+1)(J_tMJ_t - NJ_t)Q_{k-1}AP_{k-1}] \\ &\equiv J_t + p^{k-1}[N - J_tM - (r+1)(N - J_tM) + (r+1)(J_tM - N)] \\ &\equiv J_t + p^{k-1}[-(2r+1)N + (2r+1)J_tM] \equiv J_t \pmod{p^k}. \end{aligned}$$

Along with (5), this gives the desired result. Since A is similar to a unique J_t modulo p, the same is true modulo p^m.

To count the number of involutory matrices modulo p^m, we need the number $N_n(m)$ of nonsingular $n \times n$ matrices modulo p^m. This number is readily obtained, for if R is nonsingular modulo p^{k-1}, the matrices which are nonsingular modulo p^k and congruent to R modulo p^{k-1} are those matrices of the form $R + p^{k-1}S$

wherein we get incongruent matrices by taking incongruent S matrices modulo p. Thus $N_n(k)=p^{n^2}N_n(k-1)$ and hence we find that

$$N_n(m) = p^{n^2(m-1)}\cdot g_n,$$

with g_n being the well-known number of $n\times n$ nonsingular matrices modulo p.

Following the method of Hodges in [1] we let $S_0(n, t)$ be the number of distinct (incongruent) $n\times n$ matrices similar to J_t modulo p^m. As P runs through all nonsingular matrices (mod p^m), $P^{-1}J_tP$ runs through the $S_0(n, t)$ matrices similar to J_t, each matrix being repeated the same number of times as J_t. But the number of times J_t is repeated is the number of nonsingular matrices which commute with J_t, and this number is easily calculated to be

$$C_0(n, t) = N_t(m)N_{n-t}(m).$$

Thus $N_n(m)=S_0(n, t)N_t(m)N_{n-t}(m)$, so that

$$S_0(n, t) = \frac{N_n(m)}{N_t(m)N_{n-t}(m)} = \frac{p^{2t(n-t)\ (m-1)}\, g_n}{g_tg_{n-t}},$$

and hence the number N of involutory matrices modulo p^m is

$$N = \sum_{t=0}^{n} S_0(n, t) = g_n\sum_{t=0}^{n}\frac{p^{2t(n-t)(m-1)}}{g_tg_{n-t}}, \quad \text{where} \quad g_0 = 1.$$

The referee has pointed out that Theorem 1 can also be obtained from the solutions exhibited in [3].

The author would like to point out that since this article was accepted, there has appeared (in Russian) a paper which contains his Theorem 1 as a special case; however the attack is completely different. The paper is: D. A. Suprunenko, On the conjugacy of matrices over a ring of residues, Doklady Akademii Nauk, BSSR, 8 (1964)693–695. A review of this paper appears in the October 1965 issue of the Math. Reviews (R3102).

References

1. J. H. Hodges, The matrix equation $X^2-I=0$ over a finite field, this MONTHLY, 65 (1958) 518–520.

2. Jack Levine and R. R. Korfhage, Automorphisms of Abelian groups induced. by involutory matrices, general modulus, Duke Math. J., 31 (1964) 631–654.

3. Irma Reiner, The matrix congruence $X^2\equiv I$ (mod p^a), this MONTHLY, 67 (1960) 773–775.

SOME APPLICATIONS OF THE INTERIOR MAPPING PRINCIPLE

JOSEPH KIST, Pennsylvania State University

One version of the interior mapping principle asserts that a one-to-one continuous linear mapping of one Banach space upon another is bicontinuous. The closed graph theorem, which is an important consequence of the interior mapping theorem, in turn leads to the following generalization of the interior mapping principle.

THEOREM 1. *Let F be a Banach space. For each x in some nonempty set X, let F_x be a normed vector space, and u_x a continuous linear mapping of F into F_x. Assume that the family $\{u_x\}$ is separating, i.e., if f and g are distinct points of F, then there is an index x such that $u_x(f) \neq u_x(g)$. Then a linear mapping u of a Banach space E into F is continuous if and only if each of the compositions $u_x u$: $E \to F_x$ is continuous.*

Proof. The necessity is trivial. Let $\{f_n\}$ be a sequence in E such that $f_n \to f$ and $uf_n \to g$. Then

$$u_x u(f_n) \to u_x u(f), \quad \text{i.e.,} \quad u_x(uf_n) \to u_x(uf),$$

and $u_x(uf_n) \to u_x(g)$ for each x in X. Thus, $u_x(uf) = u_x(g)$ for each x in X, and since the family $\{u_x\}$ separates points of F, we have $uf = g$. By the closed graph theorem [4, p. 181], u is continuous.

That the above theorem is a generalization of the classical interior mapping principle is seen as follows. Let F and G be Banach spaces, and let v be a one-to-one continuous linear mapping of F upon G. Let u be the inverse of v. In the above theorem, replace E by the space G, $\{F_x\}$ by the singleton $\{G\}$, and $\{u_x\}$ by the singleton $\{v\}$. Since vu is the identity mapping of G, the result follows.

Theorem 1 is not new; it is essentially Theorem 7 on p. 58 of [**1**].

The object of this note is to show how the extended interior mapping principle can be used to prove several known results in the theory of Banach spaces and Banach algebras. In Section 1, it is shown that a recent result of Gil de Lamadrid is included in Theorem 1. Section 2 is concerned with a proof of the principle of uniform boundedness, and the final section is devoted to a proof of the continuity of homomorphisms between certain Banach algebras.

For some applications, it will be convenient to have Theorem 1 in the following equivalent form.

THEOREM 1′. *For each x in some nonempty index set X, let F_x be a normed vector space. Let F be a Banach space which, as a vector space, is a linear subspace of the cartesian product $P_{x \in X} F_x$ of the vector spaces F_x, and assume that the projection p_x of F into F_x is continuous for each x in X. Then a linear mapping u of a Banach space E into F is continuous if and only if the composition $p_x u$: $E \to F_x$ is continuous for each x in X.*

1. A result of Gil de Lamadrid. The following theorem is a recent result of Gil de Lamadrid [2].

THEOREM. *Let E be a Banach space, and F a normed vector space under a norm $|\ |_1$. Suppose that $|\ |_2$ is a second norm on F with respect to which F is complete and such that $k|\ |_2 \geqq |\ |_1$ for some $k > 0$. Then a linear mapping u: $E \to (F, |\ |_2)$ is continuous if and only if u: $E \to (F, |\ |_1)$ is continuous.*

Proof. In Theorem 1, let $\{F_x\}$ be the singleton $(F, |\ |_1)$, and $\{u_x\}$ the identity mapping i of $(F, |\ |_2)$ upon $(F, |\ |_1)$. By hypothesis, i is continuous. Now apply Theorem 1.

2. The principle of uniform boundedness. A popular version of the principle of uniform boundedness reads as follows.

THEOREM ON UNIFORM BOUNDEDNESS. *Let E be a Banach space. For each x in some nonempty set X, let F_x be a Banach space, and u_x a continuous linear mapping of E into F_x. If sup $\{|u_xf|:x\in X\}$ is finite for each f in E, then sup $\{|u_x|: x\in X\}$ is finite, where $|u_x|$ denotes the operator norm of u_x.*

The usual proof of this result makes direct use of the Baire category theorem for complete metric spaces. The proof given here appeals to the closed graph theorem via Theorem 1′, and is similar to a proof of one version of the principle of uniform boundedness given by Taylor [4, p. 204].

In proving the above theorem, it is convenient to make use of the notion of the normed full direct sum of the family $\{F_x: x\in X\}$ of normed vector spaces F_x. This is the set of all g in $P_{x\in X}F_x$ for which

$$|g| = \sup\{|g(x)| : x\in X\}$$

is finite, and is denoted by $\Sigma_{x\in X}F_x$. It is a linear subspace of P_xF_x, and is normed by the mapping $g\to|g|$; if each F_x is a Banach space, then so is the normed full direct sum.

To prove the principle of uniform boundedness, and with the setting as in the statement of that theorem, let $F=\Sigma_{x\in X}F_x$. Then F is a Banach space, and it is obvious that each projection p_x: $F\to F_x$ is continuous. Let u be the linear mapping of E into P_xF_x defined by $uf(x)=u_xf$, f in E, x in X. The assumption that $\sup_x|u_xf|$ is finite for each f in E means that u carries E into F. Now $p_xu(f)=uf(x)=u_xf$ for each f in E, so by assumption, $p_xu=u_x$ is continuous for each x. By Theorem 1′, there is a positive constant k such that $|uf|\leqq k|f|$ for each f in E. But

$$|uf| = \sup_x |uf(x)|, \quad \text{and so} \quad |u_xf| \leqq k|f|$$

for each x in X and f in E. Thus, $\sup_x|u_x|$ is finite.

If E and F are normed vector spaces, let $B(E, F)$ denote the normed vector space of all continuous linear mappings of E into F.

Let E and $\{F_x: x\in X\}$ be Banach spaces. If g is in $P_xB(E, F_x)$, define a linear mapping L_g of E into P_xF_x by $L_gf(x)=g(x)f$, f in E, x in X. The principle of uniform boundedness asserts that if L_g carries E into Σ_xF_x, then it is continuous, in which case g is a member of $\Sigma_xB(E, F_x)$. Moreover,

$$\begin{aligned}|L_g| &= \sup\{|L_gf| : f\in E,\ |f|\leqq 1\} = \sup_{|f|\leqq 1}\sup_x |L_gf(x)| \\ &= \sup_x\sup_{|f|\leqq 1}|g(x)f| = \sup_x|g(x)| = |g|.\end{aligned}$$

These remarks lead to the following representation theorem for continuous linear mappings between certain Banach spaces. For the case where all spaces F_x are equal to a common Banach space, this result appears in [4, p. 216].

THEOREM. *Let E and $\{F_x: x \in X\}$ be Banach spaces. For g in $\Sigma_x B(E, F_x)$, let $L_g f(x) = g(x)f$, x in X, f in E. Then the mapping $g \to L_g$ is a linear isometry of $\Sigma_x B(E, F_x)$ upon $B(E, \Sigma_x F_x)$.*

Proof. For f in E, we have

$$|L_g f| = \sup_x |L_g f(x)| \leqq |f| \sup_x |g(x)| = |f||g|,$$

so $L_g f$ is an element of $\Sigma_x F_x$. It is clear that L_g is a continuous linear mapping of E into $\Sigma_x F_x$.

It is easily seen that the mapping $g \to L_g$ is linear, and the preceding remarks show that it preserves norms; thus, it is a linear isometry. To show that this mapping is onto, let u be an element of $B(E, \Sigma_x F_x)$. Put $g(x) = p_x u$; then g is a member of $P_x B(E, F_x)$. We have

$$L_g f(x) = p_x u(f) = uf(x)$$

for x in X, and f in E. Thus, $L_g f = uf$ for each f in E. By the preceding remarks, g is a member of $\Sigma_x B(E, F_x)$. Now $L_g = u$, and this completes the proof.

3. Banach algebras. In this section, we use Theorem 1 to prove a known result from the theory of Banach algebras. For appropriate background, see [3], especially p. 75. In proving the result below, we make use of a special case of it, namely, that a homomorphism of a Banach algebra into the field C of complex numbers is continuous.

THEOREM. *A homomorphism u of a Banach algebra A into a semisimple commutative Banach algebra B is continuous.*

Proof. Let $\Sigma(B)$ be the family of all homomorphisms of B into C. Semisimplicity of B means that $\Sigma(B)$ separates points of B. For each M in $\Sigma(B)$, the composition Mu is a homomorphism of A into C. Thus, Mu is continuous for each such M, and so by Theorem 1, the homomorphism u is continuous.

Added in proof: Since this article was accepted for publication, I have found that B. E. Johnson (Proc. Cambridge Philos. Soc., 60 (1964) 171–172) has given essentially the same proof of the above result on continuity of homomorphisms. It has also come to my attention that Theorem 1′ is a special case of Theorem 1, p. 203 in A. Wilansky's *Functional Analysis*, Blaisdell, New York, 1964.

References

1. N. Dunford and J. T. Schwartz, Linear operators, Part I, Interscience Publishers, New York, 1958.

2. Jesús Gil de Lamadrid, Some simple applications of the closed graph theorem, Proc. Amer. Math. Soc., 15 (1964) 509–510.

3. C. E. Rickart, General theory of Banach algebras, Van Nostrand, Princeton, N. J., 1960.

4. A. E. Taylor, Introduction to functional analysis, Wiley, New York, 1958.

EXTREME POINTS OF THE UNIT BALL IN A SPACE OF REAL POLYNOMIALS

A. G. KONHEIM AND T. J. RIVLIN,
IBM Watson Research Center, Yorktown Heights, New York

If $p(x) = a_0 + a_1x + \cdots + a_nx^n$ and $a_n \neq 0$, then we say p has degree n. Let P_n denote the linear space of polynomials with real coefficients of degree not exceeding n. The space P_n is normed by putting

$$\|p\| = \max_{-1 \leq x \leq 1} |p(x)|. \tag{1}$$

We call a $p \in P_n$ *admissible* if $\|p\| \leq 1$, and let A_n denote the set of admissible polynomials in P_n. Then it is easy to verify that A_n is a compact convex set. We shall call $p \in A_n$ an *extremal* if p is an extreme point of A_n, i.e., if $p = \frac{1}{2}(p_1 + p_2)$ with $p_1, p_2 \in A_n$ implies $p = p_1 = p_2$. Let E_n denote the set of extremals in A_n. The purpose of this work is to characterize the (nonempty) set E_n.

Our result is contained in a theorem whose proof requires several lemmas. The proof of the first lemma is obvious and is omitted.

LEMMA 1. *If $p(x) \in E_n$ (resp. A_n) then the polynomials $-p(x)$, $p(-x)$ and $-p(-x)$ are also in E_n (resp. A_n).*

LEMMA 2. *Suppose that $p \in A_n$ and $p(x_0) = 1$, where $-1 < x_0 < 1$. Then there exists a positive integer m such that*

$$p(x) = 1 - (x - x_0)^{2m} r(x),$$

where $r(x) \geq 0$ in $[-1, 1]$ and $r(x_0) > 0$.

Proof. Since $p(x_0) = 1$

$$1 - p(x) = (x - x_0)^k r(x), \qquad k \geq 1$$

with $r(x_0) \neq 0$. If k is odd then $(x - x_0)^k$ changes sign at x_0, but $r(x)$ has one sign in some neighborhood of x_0. Hence there are points in some neighborhood of x_0 at which

$$1 - p(x) = (x - x_0)^k r(x) < 0,$$

which contradicts the assumption that $p \in A_n$. Hence $k = 2m$ and consequently $r(x) \geq 0$ in $[-1, 1]$.

LEMMA 3. *Let $p = \frac{1}{2}(p_1 + p_2)$ where $p_1, p_2 \in A_n$. Suppose that $p(x_0) = \epsilon$, $p^{(k)}(x_0) = 0$, $k = 1, 2, \cdots, k_0 - 1$, $p^{(k_0)}(x_0) \neq 0$, where $-1 \leq x_0 \leq 1$ and $\epsilon = 1$ or -1. Then for $j = 1, 2$, $p_j(x_0) = \epsilon$ and $p_j^{(k)}(x_0) = 0$, $k = 1, 2, \cdots, k_0 - 1$.*

Proof. It suffices to consider the case $\epsilon = 1$ by Lemma 1. Suppose first that $-1 < x_0 < 1$. In view of Lemma 2

$$p(x) = 1 - (x - x_0)^{2m} r(x)$$

with $r(x)\geqq 0$ in $[-1, 1]$, $r(x_0)>0$ and $k_0=2m$. Also, since $p_1, p_2\in A_n$ and

$$p = \tfrac{1}{2}(p_1 + p_2), \tag{2}$$

we must have $p_1(x_0)=p_2(x_0)=1$ and by Lemma 2, therefore,

$$p_j(x) = 1 - (x - x_0)^{2m_j}r_j(x)$$

with $r_j(x)\geqq 0$ in $[-1, 1]$ and $r_j(x_0)>0, j=1, 2$. Equation (2) implies that

$$(x - x_0)^{2m}r(x) = \tfrac{1}{2}\{(x - x_0)^{2m_1}r_1(x) + (x - x_0)^{2m_2}r_2(x)\}$$

from which we conclude that $m_j\geqq m$, $j=1$, 2. The lemma is proved for $-1<x_0<1$.

It remains to consider $x_0=\pm 1$, and it suffices (again by Lemma 1) to consider $x_0=1$. But this case is treated exactly as in the case of an interior point, except that the multiplicity k_0 of the 1 value need not be even.

We are now in a position to characterize the extremals. Let $N(p, y)$ denote the total multiplicity with which the value y is assumed by $p(x)$ in $[-1, 1]$. Let $N(p)=N(p, 1)+N(p, -1)$.

THEOREM. *If $p\in A_n$ then $p\in E_n$ if and only if $N(p)>n$.*

Proof. (i) Suppose that $N(p)>n$ and

$$p = \tfrac{1}{2}(p_1 + p_2) \tag{3}$$

with $p_1, p_2\in A_n$. In view of Lemma 3 the polynomials $p_j-p, j=1, 2$, each have zeros of total multiplicity at least $N(p)$ which together with $p, p_1, p_2\in P_n$ implies $p=p_1=p_2$, and so $p\in E_n$.

(ii) Suppose that $N(p)\leqq n$. Let $-1\leqq x_1<x_2< \cdots <x_t\leqq 1$ be all of the points in $[-1, 1]$ at which $|p(x)|=1$ and m_i the multiplicity of the zero of $p(x)-p(x_i)$ at $x=x_i$. Set

$$p_+^*(x) = \begin{cases} (-1)^{m_t} \displaystyle\prod_{\substack{i=1\\ p(x_i)=1}}^{t} (x - x_i)^{m_i} & \text{if } x_t = 1 \text{ and } p(x_t) = 1 \\ \displaystyle\prod_{\substack{i=1\\ p(x_i)=1}}^{t} (x - x_i)^{m_i} & \text{otherwise,} \end{cases} \tag{4}$$

$$p_-^*(x) = \begin{cases} (-1)^{m_t} \displaystyle\prod_{\substack{i=1\\ p(x_i)=-1}}^{t} (x - x_i)^{m_i} & \text{if } x_t = 1 \text{ and } p(x_t) = -1 \\ \displaystyle\prod_{\substack{i=1\\ p(x_i)=-1}}^{t} (x - x_i)^{m_i} & \text{otherwise,} \end{cases} \tag{5}$$

and $p^*(x)=p_+^*(x)p_-^*(x)$. Hence if $\alpha\geqq 0$ then

$$p_1(x) = p(x) + \alpha p^*(x)$$
$$p_2(x) = p(x) - \alpha p^*(x)$$

belong to A_n provided that

$$\alpha \leqq \min \left\{ \min_{-1 \leqq x \leqq 1} \frac{1 - p(x)}{p^*(x)}, \min_{-1 \leqq x \leqq 1} \frac{1 + p(x)}{p^*(x)} \right\}.$$

But $p(x) = 1 - p_+^*(x) r_+(x) = -1 + p_-^*(x) r_-(x)$, where $r_+(x) > 0$, $r_-(x) > 0$ in $[-1, 1]$. Thus

$$\min \left\{ \min_{-1 \leqq x \leqq 1} \frac{1 + p(x)}{p^*(x)}, \min_{-1 \leqq x \leqq 1} \frac{1 - p(x)}{p^*(x)} \right\} > 0$$

and hence there exists an $\alpha > 0$ such that both

$$p_1(x) = p(x) + \alpha p^*(x)$$

and

$$p_2(x) = p(x) - \alpha p^*(x)$$

belong to A_n. Since $p = \frac{1}{2}(p_1 + p_2)$ the theorem is proved.

COROLLARY. *If T_n is the Chebyshev polynomial of order n (i.e., $T_n(x) = \cos n\theta$, where $x = \cos \theta$, $0 \leqq \theta \leqq \pi$), and $n \geqq 1$ then $N(T_n, 1) = N(T_n, -1) = n$, and hence*

$$T_n \in E_k, \qquad n \leqq k \leqq 2n - 1, \qquad T_n \notin E_{2n}.$$

Note that $T_0(x) = 1 \in E_m$ for $0 \leqq m < \infty$.

REMARK 1. If $p \in A_n$ and $N(p) = 2n$ then $p = \pm T_n$. If $N(p) = 2n - 1 (n \geqq 1)$ then $p(x) = \pm T_n(ax + b)$ for some choice of a, b such that $ax + b$ maps $[-1, 1]$ into itself.

REMARK 2. As A_n is a compact subset of the $(n+1)$-dimensional linear space P_n, each element of A_n is a convex combination of at most $n + 2$ extreme points of A_n according to a well-known theorem of Carathéodory.

SOME COMBINATORIAL IDENTITIES INVOLVING LATTICE PATHS

V. K. ROHATGI, University of Alberta, Edmonton

An advanced problem proposed by Robert Breusch in this MONTHLY [1] was to find A_i ($i = 0, 1, 2, \cdots$) if for every nonnegative integer n

$$(1) \qquad \sum_{i=0}^{n} \binom{2n - 2i}{n - i} A_i = \binom{2n + 1}{n}.$$

The problem reduces to proving the identity

A REMARK ON A THEOREM OF M. A. KRASNOSELSKI

M. EDELSTEIN, Dalhousie University

If $f: K \to K$ is a continuous mapping of a closed convex subset of a normed linear space X into itself and $f(K)$ is contained in a compact subset of K then, by a well-known theorem of Schauder, there exists an $x \in K$ such that $f(x) = x$. In general no method is available for the computation of such a point. If in addition, however, f satisfies the condition

$$\|f(x) - f(y)\| \leqq \|x - y\| \qquad (x, y \in K) \tag{1}$$

and X is assumed to be a uniformly convex Banach space then Krasnoselski [2] proved that any sequence of iterates $\{F^n(x)\}$ under a "bisection" mapping $F: K \to K$ converges to a fixed point. Here F is defined by

$$F(x) = \tfrac{1}{2}(f(x) + x). \tag{2}$$

(For an exposition in English of Krasnoselski's results cf. F. F. Bonsall's Lecture Notes [1].)

It is the purpose of this note to show that the hypothesis of uniform convexity is unnecessarily stringent and, in fact, it suffices to assume that X is merely strictly convex; i.e. the unit sphere of X does not contain any line segment.

THEOREM. *Let K be a closed convex subset of a strictly convex Banach space X; $f: K \to K$ a mapping satisfying* (1), *and suppose $f(K)$ is contained in a compact subset K_1 of K. Then the sequence $\{F^n(x)\}$, where $F: K \to K$ is the mapping defined by* (2), *converges to a fixed point for all $x \in K$.*

Proof. The strict convexity of X combined with (1) imply that

$$\|F(x) - y\| < \|x - y\| \tag{3}$$

whenever $f(y) = y$ and $f(x) \neq x$.

Indeed the open line segment $]x, f(x)[$ is contained in the open ball $\{w \mid \|w - y\| < \|x - y\|\}$ by the strict convexity of X; in particular its midpoint $F(x)$ is in this open ball as asserted by (3). Also, since $\{F^n(x)\}$ is a subset of the closed convex hull of $K_1 \cup \{x\}$, which is compact by the Mazur Theorem [3], we know that $\{F^n(x)\}$ contains a subsequence $\{F^{n_i}(x)\}$ which converges to some $p \in K$.

To prove our theorem it obviously suffices to show that $f(p) = p$. (That $\lim F^n(x) = p$ follows then immediately from (1).) Suppose that this is not so. Then no element of $\{F^n(x)\}$ can be fixed under f; for $f(F^k(x)) = F^k(x)$ implies $F^{k+i}(x) = F^k(x)$, $i = 1, 2, \cdots$, and the whole sequence $\{F^n(x)\}$ converges to a fixed point (of necessity p).

Thus (3) applies to all pairs $(F^n(x), y)$ $n = 1, 2, \cdots$, and we have

$$\|F^{n+1}(x) - y\| < \|F^n(x) - y\| \qquad (n = 1, 2, \cdots). \tag{4}$$

Let $B(F(p), r) = \{w \mid \|w - F(p)\| < r\}$ with $r = \frac{1}{2}(\|p-y\| - \|F(p)-y\|)$. F being obviously continuous, an open ball B_1 centered at p, and of radius $r_1 \leqq r$, exists so that $F(B_1) \subset B$. If now $F^k(x) \in B_1$ for some k then $F^{k+1}(x) \in B$. Using (4) we now get $\|F^{k+i}(x) - y\| < \|F(p) - y\| + r = \frac{1}{2}(\|F(p)-y\| + \|p-y\|)$. Hence,

$$\begin{aligned} \|F^{k+i}(x) - p\| &= \|(F^{k+i}(x) - y) + (y - p)\| \\ &\geqq \|y - p\| - \|F^{k+i}(x) - y\| \\ &> \|y - p\| - \tfrac{1}{2}(\|F(p) - y\| + \|p - y\|) \\ &= \tfrac{1}{2}(\|p - y\| - \|F(p) - y\|) = r, \end{aligned}$$

which is clearly incompatible with $\lim F^{n_i}(x) = p$. Thus $f(p) = p$ and our theorem is established.

References

1. F. F. Bonsall, Lectures on some fixed point theorems of functional analysis, Tata Institute of Fundamental Research, Bombay, 1962.

2. M. A. Krasnoselski, Two remarks on the method of successive approximations, Uspehi Math. Nauk (N.S.), 10, No. 1, 63 (1955) 123–127, (Russian).

3. S. Mazur, Über die kleinste konvexe Menge, die eine gegebene kompakte Menge enthält, Studia Math., 2 (1930) 7–9.

ON HIGHLY COMPOSITE CONSECUTIVE INTEGERS

P. K. Subramanian, Miami University and
S. F. Becker, Rutgers University

It is well known that for any positive integer k, one can find k consecutive numbers which are all composite. The following question arose out of a problem which appeared in this Monthly [1].

> "Is it possible to find sequences of consecutive integers each member of which is as 'highly' composite as desired?"

The answer is in the affirmative. We state and prove the following:

Theorem. *For any positive integers n and r, and any sequence $\{p_i\}_{i=1}^{i=r}$ of positive integers, there exists a sequence $\{k+i\}_{i=1}^{i=n}$ such that for every integer $(k+i)$ there exists a sequence $\{c_{ij}\}_{j=1}^{j=r}$ where*

$$k + i \equiv 0 \mod \prod_{j=1}^{r} c_{ij}^{p_j} \qquad 1 \leqq i \leqq n$$

and $(c_{ij}, c_{ik}) = 1$ for $j \neq k$. Here $j, k = 1, 2, \cdots, r$.

In other words, suppose a positive integer n and any arbitrary finite sequence $p_1, p_2, \cdots, p_r$ of positive integers are given. Then, there exist infinitely many sequences of n consecutive integers, where each integer has as factors (besides, possibly others) r relatively prime numbers raised to any preassigned powers. These preassigned powers are members of the sequence $p_1, p_2, \cdots, p_r$.

We construct a sequence $\{A_i\}$ of relatively prime integers and prove some properties of this sequence, that will be needed in the proof of the main theorem.

Let q be any positive integer greater than unity. Consider the sequence $\{A_i\}$ where

$$A_1 = q - 1$$
$$A_2 = (2A_1 + 1)^2$$
$$\cdots\cdots$$
$$A_{n+1} = \left(2\prod_{i=1}^{n} A_i + 1\right)^2.$$
$$\cdots\cdots\cdots\cdots$$

We next prove four lemmas, three of which concern the above sequence.

LEMMA 1. *For every* $i \neq j$, $(A_i, A_j) = 1$.

Proof. Without loss of generality we may assume that $A_i < A_j$.

$$A_j = \left(2\prod_{k=1}^{j-1} A_k + 1\right)^2 \quad \text{by construction}$$
$$= \left(2\prod_{k=1}^{j-1} A_k\right)^2 + 4\prod_{k=1}^{j-1} A_k + 1$$

which implies $A_j \equiv 1 \bmod A_i$.

LEMMA 2. *For every positive integer* n, $\prod_{i=1}^{n} A_i + 1 \equiv 0 \bmod q$.

Proof. We prove by induction. Obviously when $n = 1$

$$\prod_{i=1}^{1} A_i + 1 = (q - 1) + 1 \equiv 0 \bmod q.$$

Assume that for $n = l$, $\prod_{i=1}^{l} A_i + 1 \equiv 0 \bmod q$; i.e. $\prod_{i=1}^{l} A_i = kq - 1$ for some k. Now

$$\prod_{i=1}^{l+1} A_i = A_{l+1}\prod_{i=1}^{l} A_i = \left(2\prod_{i=1}^{l} A_i + 1\right)^2 \prod_{i=1}^{l} A_i = [2(kq - 1) + 1]^2(kq - 1)$$

which implies $\prod_{j=1}^{l+1} A_i \equiv -1 \bmod q$, and the lemma is proved.

LEMMA 3. *For every positive integer* n, $A_{n+1} \equiv 1 \bmod q$.

Proof.

$$A_{n+1} = \left[2\prod_{i=1}^{n} A_i + 1\right]^2 = [2(kq - 1) + 1]^2$$

for some k by Lemma 2. This completes the proof.

LEMMA 4. *For all positive integers* $s, n, p_1, p_2, \cdots, p_r$ *and every positive integer* $q(>1)$, *there exist positive integers* $b, B_1, \cdots, B_r$ *such that* (i) $(B_i, B_j)=1$ *for* $i\neq j$ *and* (ii) $bq-s\equiv 0 \bmod \prod_{i=1}^r B_i^{p_i}$.

Proof. For $B_1, B_2, \cdots, B_r$ we may choose the first r members of the sequence $\{A_i\}$ constructed above. (i) is obvious by Lemma 1. We now choose

$$b = s\,\frac{\left[\prod_{i=1}^{r} A_i^{p_i} + 1\right]}{q}. \tag{1}$$

Since $A_1=q-1\equiv -1 \bmod q$, it follows that $A_1^{p_1}\equiv 1 \bmod q$ if p_1 is even and $A_1^{p_1}\equiv -1 \bmod q$ if p_1 is odd.

Case I. Let p_1 be odd. By Lemma 3, for all $i\geqq 2$,

$$A_i \equiv 1 \bmod q \Rightarrow A_i^{p_i} \equiv 1 \bmod q.$$

Hence

$$\prod_{i=1}^{r} A_i^{p_i} \equiv -1 \bmod q.$$

This implies $q\mid \prod_{i=1}^r A_i^{p_i}+1$, which shows that b is an integer.

Substituting the value of b from (1) in $bq-s$ we have

$$bq - s = s\left[\prod_{i=1}^{r} A_i^{p_i}\right]$$

which implies $bq-s\equiv 0 \bmod \prod_{i=1}^r A_i^{p_i}$.

Case II. Let p_1 be even. Since $p_1' = p_1+1$ is odd we may apply Case I to p_1'. Since, if $A_1^{p_1+1}\mid bq-s$ then so does $A_1^{p_1}$, then this proves Case II. Thus (ii) holds and the lemma is proved.

Proof of the main theorem. Let r be given and $p_1, p_2, \cdots, p_r$ be the corresponding sequence stated in the theorem. We use induction.

The theorem is true for $n=1$, since we may choose $2^{p_1}3^{p_2}5^{p_3}\cdots t_r^{p_r}$ for the required integer where t_r is the rth prime.

Assume then, that the theorem is true for $n=l$; that is, there exist consecutive integers

$$k+l,\ k+l-1,\ \cdots,\ k+1, \text{ such that } k+i\equiv 0 \bmod \prod_{j=1}^{r} c_{ij}^{p_j}, \qquad 1\leqq i\leqq l \tag{2}$$

for some c_{ij}'s such that $(c_{ij}, c_{ik})=1$ for $k\neq j$. Now, (2) implies that the sequences

$$m(k+l)! - (k+l),\ m(k+l)! - (k+l-1),\ \cdots,\ m(k+l)! - (k+1); \qquad m = 1, 2, \cdots \tag{3}$$

also have the required property, since each member of the sequence (2) divides

the corresponding member of the sequences (3) for every m.

If we set

$$q = (k + l)!,$$
$$s = k \quad \text{and}$$
$$B_i = A_i \quad (1 \leqq i \leqq r)$$

in Lemma 4, then there exists a positive integer m' such that

$$m'(k + l)! - k \equiv 0 \mod \prod_{i=1}^{r} A_i^{p_i}.$$

By setting $m = m'$ in (3) we have constructed a sequence of $(l+1)$ consecutive integers

$$m'(k + l)! - k, m'(k + l)! - (k + 1), \cdots, m'(k + l)! - (k + l)$$

having the required property. This completes the induction.

We have now proved that for every positive integer n there exist n consecutive integers with the properties stated in the theorem.

For any such sequence $k'+1, k'+2, \cdots, k'+n$, the sequences

$$m(k' + n)! - (k' + n), m(k' + n)! - (k' + n - 1), \cdots, m(k' + n)! - (k' + 1); \quad m = 1, 2, \cdots,$$

also have the desired properties. Thus there are infinitely many such sequences. This proves the theorem.

In proving the theorem we have not only shown the existence of consecutive integers with the properties required by the theorem but have also indicated a recursive algorithm for constructing such sequences.

The authors gratefully acknowledge the suggestions and generous encouragement of Dr. Richard F. DeMar.

The first author is now at Ohio State University, Columbus, Ohio.

Reference

1. Problem No. E 1709, this MONTHLY, 71 (1964) 680.

FUNCTION SPACES OF INVERTIBLE SPACES

S. A. NAIMPALLY, Michigan State University

1. Introduction. Let S be a Hausdorff space and F the function space of all continuous self-mappings on S with the compact-open topology ([4] p. 221). A question naturally arises, "if S is invertible ([2]) is F invertible?" Although this question is still not resolved, we propose to present in this note a few partial answers in that direction.

We recall a few basic definitions.

1.1. *S is an invertible space if for each nonempty open subset U of S, there exists a self-homeomorphism h of S such that $h(S-U)\subset U$* [2].

1.2. *S is near-homogeneous at $p\in S$ if for each open neighborhood U of p and each point $x\in S$, there exists a self-homeomorphism of S such that $h(x)\in U$* [1].

If in the above definitions U is an element of a sub-basis of S, then S will be called subinvertible and sub-near-homogeneous respectively. We first prove a useful lemma.

1.3. LEMMA. *If h is a self-homeomorphism of a topological space S, then the induced self-map h^* of F, defined by $h^*(f)=h\circ f$ for $f\in F$, is a homeomorphism.*

Proof. First, h^* is well-defined and since $h^*(f)=h^*(g)$ for f, $g\in F$ implies $f=g$, h^* is one to one. In order to prove that h^* is continuous, consider a sub-basis element

$$W = \{f \mid f\in F, f(C)\subset U\}\subset F,$$

where C is a compact and U an open subset of S. Now $h^{*-1}(W)=\{g\mid g\in F, g(C)\subset h^{-1}(U)\}$ which is open in S since $h^{-1}(U)$ is open in S. This shows that h^* is continuous. We can similarly show that h^{*-1} is continuous which proves that h^* is a self-homeomorphism of F.

2. Properties of F. Let S be an invertible Hausdorff space.

2.1. THEOREM. *F is sub-near-homogeneous at each of its points.*

Proof. Let W be a sub-basis element of S as in the proof of Lemma 1.3. If $g\in F$ is an element of W the theorem is obvious. If $g\notin W$ then $g(C)\not\subset U$. But $g(C)$ is a compact and (since S is Hausdorff) closed subset of S. By a theorem of [2], there exists a self-homeomorphism h of S such that $h[g(C)]\subset U$. By Lemma 1.3, h^* is a self-homeomorphism of F and $h^*(g)\in W$.

Question. Is F sub-invertible?

Let $M\subset F$ be the subset consisting of all constant maps. If an element of M maps the whole of S onto a point $p\in S$ we denote this map by fp. We now prove some properties of M.

2.2. THEOREM. *M is closed in F.*

Proof. Let $g\in F-M$. Then $g[S]$ consists of at least two distinct points x and y. Since S is Hausdorff, there are two disjoint open sets U and V containing x and y respectively. The function g belongs to the open set

$$\{f\mid f(x)\in U\}\cap\{f\mid f(y)\in V\}$$

which clearly does not meet M. This shows that $F-M$ is open and so M is closed.

The following lemma is obvious.

2.3. LEMMA. *If h is a self-homeomorphism of S, then $h^*(M)=M$, where h^* is as in* 1.3.

2.4. THEOREM. *F is near-homogeneous at each point of M.*

Proof. Let $fp \in M$ and let $\bigcap_{i=1}^{n} (C_i, U_i)$ be a basis element contained in any given open neighborhood of fp, where $(C_i, U_i) = \{g \mid g \in F,\ g(C_i) \subset U_i,\ C_i,\ U_i$ being compact and open subsets of F respectively$\}$. Clearly the open set $U = \bigcap_{i=1}^{n} U_i$ contains p and if $g \in F$ then $C = \bigcup_{i=1}^{n} g(C_i)$ is compact and so closed in S. Since S is invertible there exists a self-homeomorphism h of S such that $h(C) \subset U$. Then $h^*(g) \in \bigcap_{i=1}^{n} (C_i, U_i)$ and so F is near-homogeneous at fp.

Conjecture. If $f \in F - M$ then F is not near-homogeneous at f.

*Acknowledgment.*The author wishes to thank Professors Doyle and Hocking for introducing him to this subject and for support from NSF grant GP-31.

The author is at present at University of Alberta, Edmonton, Canada.

References

1. C. E. Burgess, Certain types of homogeneous continua, Proc. Amer. Math. Soc., 6 (1955) 348–350.
2. P. H. Doyle and J. G. Hocking, Invertible spaces, this MONTHLY, 68 (1961) 959–965.
3. J. G. Hocking and G. S. Young, Topology, Addison-Wesley, Reading, Mass., 1961.
4. J. L. Kelley, General Topology, Van Nostrand, Princeton, N. J., 1955.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

A SIMPLE EXAMPLE OF A WEIERSTRASS FUNCTION

A. A. BLANK, New York University

The subject of nowhere differentiable continuous functions or Weierstrass functions has received a great deal of attention (see Singh [1]). In consideration of the relative abundance of such functions there is a remarkable dearth of examples which are both simple in conception and simple in the verification of their properties. A simple numerical example was given by Swift [2]. Here we give a geometrical example which demands no sophistication either in its construction or in its demonstration, and is easily visualized.

We shall define a Weierstrass function as the limit of a sequence of polygonal approximations. Let (x_1, y_1) and (x_2, y_2) be consecutive vertices of an approximating polygon where $x_1 < x_2$ and $y_2 \neq y_1$. Let $h = \frac{1}{3}(x_2 - x_1)$ and let $k = \lambda(y_2 - y_1)$ for a fixed positive λ. To define the next approximating polygon we shall replace the segment joining (x_1, y_1) to (x_2, y_2) by a zigzag (Figure 1) joining the consecutive vertices

$$(x_1, y_1),\ (x_1 + h, y_2 - k),\ (x_2 - h, y_1 + k),\ (x_2, y_2).$$

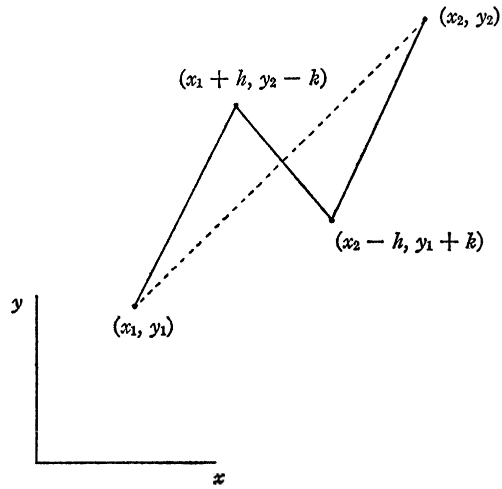


FIG. 1

The differences in ordinate of the successive vertices are

$$(1-\lambda)(y_2-y_1),\ (2\lambda-1)(y_2-y_1),\ (1-\lambda)(y_2-y_1).$$

We choose λ so that the greatest absolute difference in the successive ordinates is $\mu|y_2-y_1|$, where

$$\mu = 1-\lambda < 1. \tag{1}$$

As we shall see, this decrease in the absolute difference yields continuity in the limit.

We denote the slope of the segment joining (x_1, y_1) to (x_2, y_2) by m and obtain for the slopes of the three consecutive segments of the zigzag, in order,

$$3(1-\lambda)m,\ 3(2\lambda-1)m,\ 3(1-\lambda)m.$$

We choose λ so that the minimum steepness of absolute slope of the three segments is $\nu|m|$ where

$$\nu = 3(1-2\lambda) > 1. \tag{2}$$

As we shall see, this increase in steepness yields non-differentiability in the limit.

To satisfy (1) and (2) we choose any λ for which $0<\lambda<1/3$. Observe that under conditions (1) and (2) the entire zigzag on the open interval (x_1, x_2) remains between the lines $y=y_1$, $y=y_2$.

For the construction of a Weierstrass function by successive approximation we fix $\lambda<1/3$. We begin with the segment of the line $y=x$ for $0\leqq x\leqq 1$. The first approximating polygon is the zigzag constructed by the above method. The second approximating polygon is obtained from the first by using the zigzag

construction on each third, $i/3 \leqq x \leqq (i+1)/3$, $(i=0, 1, 2)$, of the interval. We iterate the scheme by applying the zigzag construction to each segment of an approximating polygon in order to obtain the next approximation. Figure 2 shows the fourth approximating polygon for $\lambda=0.3$. The abscissae of successive vertices of the nth approximating polygon are the ternary points $x_{n,i}=i/3^n$, $(i=0, 1, 2, \cdots, 3^n)$ and since these vertices are vertices of all succeeding polygons, they will be points on the graph of the prospective Weierstrass function.

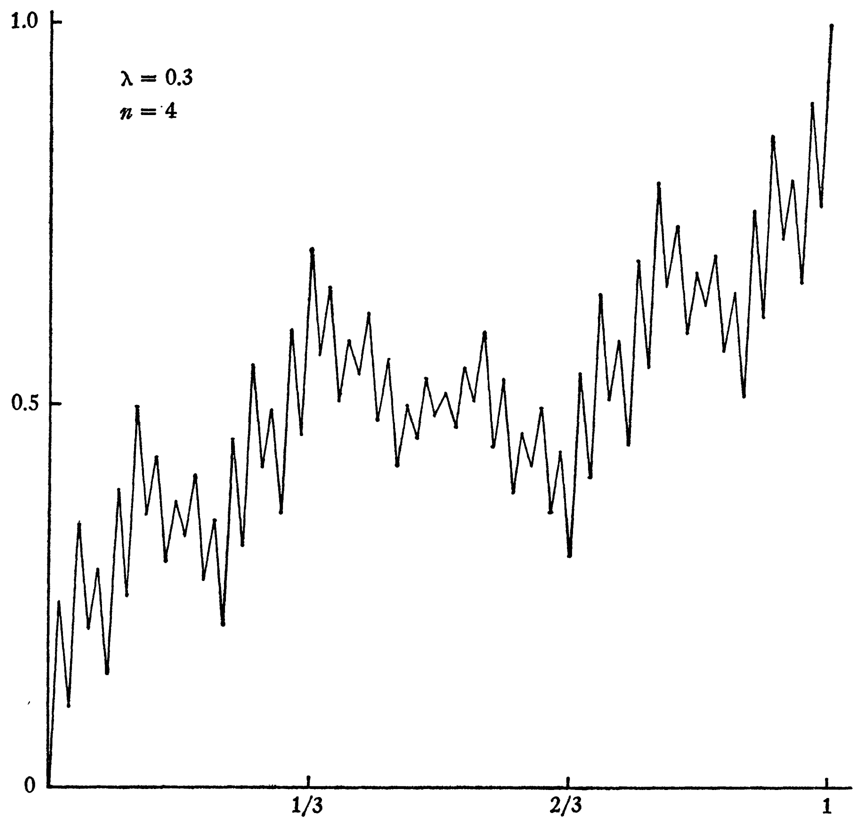


FIG. 2

This sequential construction yields a function f defined on all ternary points $x_{n,i}$ and we shall define the Weierstrass function as the continuous extension of f to all points of the interval $[0, 1]$.

First, we observe that our method of construction yields for the difference in function values at successive ternary points of order n

$$|f(x_{n,i}) - f(x_{n,i-1})| \leqq \mu^n, \qquad (i = 1, 2, \cdots, 3^n), \tag{3}$$

where μ is given by (1). For the slope of the segment joining the corresponding vertices, the construction yields

$$\frac{|f(x_{n,i}) - f(x_{n,i-1})|}{x_{n,i} - x_{n,i-1}} \geqq \nu^n, \qquad (i = 1, 2, \cdots, 3^n), \tag{4}$$

where ν is given by (2).

Given any real number r in the open interval (0, 1), we set $a_n = x_{n,i-1}$, $b_n = x_{n,i}$ where i is the unique index defined by

$$x_{n,i-1} \leqq r < x_{n,i}.$$

Clearly $a_{n+1} \geqq a_n$ and $b_{n+1} \leqq b_n$. As we have already observed, $f(a_{n+1})$ and $f(b_{n+1})$ both lie in the closed interval $[f(a_n), f(b_n)]$. (The use of this notation for the interval is not meant to imply $f(a_n) < f(b_n)$.) From (3) it follows that the nested set of intervals $[f(a_n), f(b_n)]$ contains precisely one real number, and this is taken as the value $f(r)$.

Given a positive ϵ we choose n so large that $2\mu^n < \epsilon$ and we take $\delta = 4^{-n}$. If $|x-r| < \delta$ and $x < r$ then, by (3),

$$|f(x) - f(r)| \leqq |f(x) - f(a_n)| + |f(a_n) - f(r)| \leqq 2\mu^n < \epsilon.$$

Similarly, if $|x-r| < \delta$ and $x > r$,

$$|f(x) - f(r)| \leqq |f(x) - f(b_n)| + |f(b_n) - f(r)| < \epsilon.$$

Thus continuity is established.

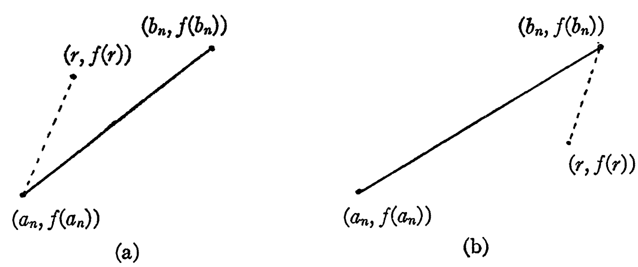


FIG. 3

Nondifferentiability is established with equal ease. First, if r is a ternary point of order n, $r = i/3^n$, then r is a ternary point of every higher order: $r = i3^k/3^{n+k}$. On any neighborhood of r it follows, from (4), that no bound can be placed on the steepness of the chords of the graph of f with one endpoint at $(r, f(r))$. Thus f is not differentiable at any ternary point. Next, if r is not a ternary point, we have strict inequality

$$a_n < r < b_n$$

for all n. It is geometrically clear that the steepness of one of the two chords join-

ing $(r, f(r))$ to the points $(a_n, f(a_n))$ and $(b_n, f(b_n))$ is at least as great as the steepness of the chord C joining $(a_n, f(a_n))$ to $(b_n, f(b_n))$. If, for example, $f(a_n) < f(b_n)$ then, if $(r, f(r))$ lies above C the chord to $(a_n, f(a_n))$ is steeper (Fig. 3a), if $(r, f(r))$ lies on C the chords are equally steep, and if $(r, f(r))$ lies below C the chord to $(b_n, f(b_n))$ is steeper (Fig. 3b). A similar argument holds if $f(a_n) > f(b_n)$. Analytically, we can treat all cases simultaneously as follows. We have

$$\frac{|f(b_n) - f(r)|}{b_n - r} + \frac{|f(r) - a_n|}{r - a_n} > \frac{|f(b_n) - f(r)| + |f(r) - f(a_n)|}{b_n - a_n}$$
$$\geqq \frac{|f(b_n) - f(a_n)|}{b_n - a_n} \geqq \nu^n,$$

where the last line follows from (2). It follows that at least one of the slopes

$$\frac{f(b) - f(r)}{b_n - r} \quad \text{or} \quad \frac{f(r) - f(a_n)}{r - a_n}$$

is absolutely greater than $\frac{1}{2}\nu^n$. Again we see that no bound can be placed on the steepness of the chords with one endpoint at $(r, f(r))$ in any neighborhood of r. We conclude that f is nowhere differentiable.

References

1. A. N. Singh, The Theory and Construction of Non-Differentiable Functions, Lucknow University Studies, No. 1, Faculty of Science, Lucknow University, 1935; (republished by Chelsea, New York, in Squaring the Circle and Other Monographs).

2. W. C. Swift, Simple constructions of non-differentiable functions and space-filling curves, this MONTHLY, 68 (1961) 653–655.

ON DEPENDENCE IN ABELIAN GROUPS

H. SCHWERDTFEGER, McGill University, Montreal

Let G be an additive abelian group without torsion and $x, y, x_1, x_2, \cdots, y_1, y_2, \cdots$ its elements. Small Greek letters denote rational integers; m, n, r are natural numbers.

We begin with a few well-known definitions: An element x is said to depend on (the elements of) a set $X = (x_1, \cdots, x_m) \subset G$ if for suitable integral coefficients $\alpha \neq 0, \alpha_\mu$:

$$\alpha x = \sum_{\mu=1}^{m} \alpha_\mu x_\mu; \qquad \text{in symbols: } x \vdash X.$$

Also $X \vdash Y = (y_1, \cdots, y_n)$ if for suitable $\alpha_\mu \neq 0$

$$\alpha_\mu x_\mu = \sum_{\nu=1}^{n} \beta_{\mu\nu} y_\nu \qquad (\mu = 1, 2, \cdots, m). \tag{1}$$

The set X (or its elements) is said to be independent if from any relation $\sum_{\mu=1}^{m}\alpha_\mu x_\mu=0$ follows that $\alpha_1=0, \cdots, \alpha_m=0$; otherwise X is dependent.

LEMMA 1. *If X depends on Y: $X\vdash Y$, and $Y\vdash Z$, where $Z=(z_1 \cdots, z_r)\subset G$, then $X\vdash Z$.*

REMARK. It appears that this transitivity is usually taken for granted (cf. P. et M. L. Dubreil, Leçons d'algèbre moderne, 2nd ed. p. 279; L. Fuchs, Abelian Groups, p. 32–33; A. G. Kurosh, The Theory of Groups, p. 139); since it is not entirely trivial, a proof is given in this note.

Proof. According to the supposition we have (1) and

$$\beta_\nu y_\nu = \sum_{\rho=1}^{r}\gamma_{\nu\rho}z_\rho, \qquad \beta_\nu \neq 0, \qquad (\nu = 1, 2, \cdots, n). \tag{2}$$

Let $\beta=$L.C.M. $[\beta_1, \cdots, \beta_n]=\beta'\beta_1=\beta''\beta_2= \cdots =\beta^{(n)}\beta_n$. Then by (1) and (2)

$$\beta\alpha_\mu x_\mu = \sum_{\nu=1}^{n}\beta_{\mu\nu}\beta^{(\nu)}\cdot\beta_\nu y_\nu = \sum_{\nu=1}^{n}\sum_{\rho=1}^{r}\beta_{\mu\nu}\beta^{(\nu)}\cdot\gamma_{\nu\rho}z_\rho = \sum_{\rho=1}^{r}\delta_{\mu\rho}z_\rho.$$

Thus, since $\beta\alpha_\mu\neq 0$, it follows that $X\vdash Z$.

DEFINITION. *X is equivalent to Y, in symbols: $X \dashv\vdash Y$, if*

$$X \vdash Y \quad \text{and} \quad Y \vdash X.$$

By Lemma 1 equivalence is transitive: If $X\dashv\vdash Y$ and $Y\dashv\vdash Z$ then $X\dashv\vdash Z$.

LEMMA 2. *If X is independent and $X\vdash Y$, then $m\leqq n$.*

Proof indirect. Assume that $n\leqq m-1$ and therefore

$$\alpha_\mu x_\mu = \sum_{\nu=1}^{m-1}\beta_{\mu\nu}y_\nu, \qquad \alpha_\mu \neq 0 \qquad (\mu = 1, \cdots, m).$$

Integers $\xi_1, \cdots, \xi_m$, not all zero, can be found so that $\sum_{\mu=1}^{m}\xi_\mu\beta_{\mu\nu}=0$ (as solution of $m-1$ linear homogeneous equations in m unknowns) and hence

$$\sum_{\mu=1}^{m}\xi_\mu\alpha_\mu x_\mu = 0.$$

Since the x_μ are independent we conclude $\xi_\mu\alpha_\mu=0$, in contradiction to the assumption.

Now follows easily the so-called

STEINITZ REPLACEMENT THEOREM. *Let $X\subset G$ be independent and $X\vdash Y$. Then $m\leqq n$ and m of the elements $y_1, \cdots, y_n$ of Y can be replaced by the m elements $x_1, \cdots, x_m$ of X whereby Y is changed into a set $Y^{(m)}\dashv\vdash Y$.*

Proof by induction. If $m=1$, $x_1\neq 0$ and $\alpha x_1=\sum_{\nu=1}^{n}\beta_\nu y_\nu$, $\alpha\neq 0$, as $n\geqq 1$, at

least one $\beta_\nu y_\nu \neq 0$; say $\beta_1 y_1 \neq 0$. Then

$$\beta_1 y_1 = \alpha_1 x_1 - \sum_{\nu=2}^{n} \beta_\nu y_\nu;$$

thus y_1, and hence $Y \vdash (x_1, y_2, \cdots, y_n)$.

By Lemma 2, $n \geqq m$ and we may assume that in (1) $\beta_{mm} y_m \neq 0$. For the induction we suppose that

$$Y^{(m-1)} = (x_1, \cdots, x_{m-1}, y_m, \cdots, y_n) \vdash\!\dashv Y. \tag{3}$$

By (1), $\beta_{mm} y_m = -\beta_{m1} y_1 - \cdots - \beta_{mm-1} y_{m-1} + \alpha_m x_m - \beta_{mm+1} y_{m+1} - \cdots - \beta_{mn} y_n$ where $\alpha_m \neq 0$ so that

$$y_m \vdash (y_1, \cdots, y_{m-1}, x_m, y_{m+1}, \cdots, y_n) = Y'$$

and x_m really occurs. Further, by (3), $(y_1, \cdots, y_{m-1}) \vdash Y^{(m-1)}$; therefore

$$Y' \vdash (x_1, \cdots, x_m, y_m, y_{m+1}, \cdots, y_n) = Y''$$

and also $y_m \vdash Y''$ other than by the trivial relation $y_m = y_m$ whence

$$y_m \vdash (x_1, \cdots, x_m, x_{m+1}, \cdots, y_n) = Y^{(m)}.$$

Thus $Y' \vdash Y'' \vdash Y^{(m)}$ and so, by Lemma 1, $Y' \vdash Y^{(m)}$. On the other hand $Y \vdash Y^{(m-1)} \vdash Y'$ so that $Y \vdash Y'$. Therefore $Y \vdash Y^{(m)}$. Since obviously $Y^{(m)} \vdash Y$, the theorem is now proved.

SOME EXTENSIONS OF THE INTEGRAL TEST

G. T. Cargo Syracuse University

1. Introduction. It seems natural to inquire whether or not, in the Maclaurin-Cauchy integral test for the convergence of an infinite series, it is sufficient to assume that the function in question is of bounded variation instead of positive and decreasing. Some fifty-four years ago, G. H. Hardy [3, p. 127] proved that such a condition is, indeed, sufficient provided that the function possesses a continuous derivative. The main purpose of this note is to eliminate Hardy's smoothness requirement, but we shall proceed toward this goal at a rather leisurely pace in order to acquire some historical and mathematical insight.

2. Background. For the sake of completeness, we now prove Hardy's theorem (in a somewhat modified form) as well as a closely related result due to R. W. Brink.

Theorem 1 (Hardy). *Let f be a nonnegative function which is defined and has a continuous derivative on $[0, \infty)$. Then $\sum_{k=1}^{\infty} f(k)$ and $\int_0^\infty f(t)dt$ converge or diverge together provided*

$$\int_0^\infty |f'(t)|\, dt < \infty. \tag{1}$$

Proof. Denoting the greatest integer not exceeding t by $[t]$ and using elementary properties of the Riemann-Stieltjes integral, we obtain, for each positive integer n,

$$\sum_{k=1}^{n} f(k) - \int_0^n f(t)\,dt = \int_0^n f(t)\,d[t] - \int_0^n f(t)\,dt$$
$$= \int_0^n f(t)\,d\{[t] - t\}$$
$$= -\int_0^n \{[t] - t\}\,df(t)$$
$$= \int_0^n \{t - [t]\}f'(t)\,dt.$$

Consequently,

$$\left| \sum_{k=1}^{n} f(k) - \int_0^n f(t)\,dt \right| \leqq \int_0^n |f'(t)|\,dt \qquad (n = 1, 2, \cdots), \tag{2}$$

which, in virtue of (1), completes the proof. We note, incidentally, that (2) is closely connected with the Euler summation formula (cf. [1, pp. 201–202]).

In 1919 R. W. Brink [2, p. 42] observed that Hardy's theorem is an immediate consequence of the following simple result.

THEOREM 2 (BRINK). *Let f be a nonnegative function defined on $[0, \infty)$ which is Riemann integrable over the interval $[0, r]$ for each positive number r. Then $\sum_{k=1}^{\infty} f(k)$ and $\int_0^\infty f(t)dt$ converge or diverge together provided*

$$\sum_{k=1}^{\infty} \sup\{ |f(k) - f(t)| : k - 1 \leqq t \leqq k\} < \infty. \tag{3}$$

Proof. Since

$$\begin{aligned} \left| \sum_{k=1}^{n} f(k) - \int_0^n f(t)\,dt \right| &= \left| \sum_{k=1}^{n} \int_{k-1}^{k} \{f(k) - f(t)\}\,dt \right| \\ &\leqq \sum_{k=1}^{n} \sup\{ |f(k) - f(t)| : k - 1 \leqq t \leqq k\} \end{aligned} \tag{4}$$

holds for each positive integer n, the desired conclusion follows at once from (3).

The interested reader should encounter no difficulties in proving that: (a) Theorem 2 implies Theorem 1; (b) Theorems 1 and 2 are, indeed, generalizations of the integral test.

3. An intermediate result. If one is willing to retain the assumption that the function is nonnegative, then it is easy to establish the desired extension of Hardy's theorem.

THEOREM 3. *Let f be a nonnegative function defined on* $[0, \infty)$. *Then* $\sum_{k=1}^{\infty} f(k)$ *and* $\int_0^{\infty} f(t)dt$ *converge or diverge together provided*

$$\sup\{V_0^n f: n = 1, 2, \cdots\} < \infty, \tag{5}$$

where $V_0^n f$ *denotes the total variation of f on* $[0, n]$.

Proof. Let n be a fixed positive integer, let ϵ be a positive real number, and, for each integer k satisfying $1 \leqq k \leqq n$, select a point t_k in the interval $[k-1, k]$ in such a way that $\sup\{|f(k)-f(t)| : k-1 \leqq t \leqq k\} < \epsilon/n + |f(k)-f(t_k)|$. Then it follows from (4) that

$$\begin{aligned} \left|\sum_{k=1}^{n} f(k) - \int_0^n f(t)\,dt\right| &< \epsilon + \sum_{k=1}^{n} |f(k) - f(t_k)| \\ &\leqq \epsilon + \sum_{k=1}^{n} \{|f(t_k) - f(k-1)| + |f(k) - f(t_k)|\} \\ &\leqq \epsilon + V_0^n f; \end{aligned}$$

since this inequality holds for each positive number ϵ, we conclude that

$$\left|\sum_{k=1}^{n} f(k) - \int_0^n f(t)\,dt\right| \leqq V_0^n f \qquad (n = 1, 2, \cdots), \tag{6}$$

from which the desired result follows at once.

If in Theorem 3 we also assume that f is continuous, then (6) follows from the fact that, for each positive integer n,

$$\left|\sum_{k=1}^{n} f(k) - \int_0^n f(t)\,dt\right| = \left|-\int_0^n \{[t] - t\}\,df(t)\right| \leqq V_0^n f.$$

We now present still another proof of Theorem 3 which is connected with numerical analysis and may be of some interest to the reader.

If a function f is Riemann integrable on a finite interval $[a, b]$, then the sequence of Riemann sums

$$\frac{b-a}{n} \sum_{k=1}^{n} f\left(a + k\,\frac{b-a}{n}\right) \qquad (n = 1, 2, \cdots)$$

converges to the limit $\int_a^b f(x)dx$ as n approaches infinity. It is meaningful to ask what are sufficient conditions on f in order that the sequence of Riemann sums converge with a certain specified rapidity (cf. [6, pp. 36–37]). That this question is related to extensions of the integral test is perhaps best seen by means of the following result due to Pólya (cf. [6, p. 37]): *If a function g has finite total variation V on* $[0, 1]$, *then*

$$\left|\int_0^1 g(x)\,dx - (1/n)\sum_{k=1}^{n} g(k/n)\right| \leqq V/n, \qquad (n = 1, 2, \cdots).$$

Now, if f is a function satisfying condition (5), then

$$\int_0^n f(t)\,dt - \sum_{k=1}^n f(k) = n\int_0^1 f(nx)\,dx - \sum_{k=1}^n f(k)$$
$$= n\left[\int_0^1 g_n(x)\,dx - (1/n)\sum_{k=1}^n g_n(k/n)\right],$$

where $g_n(x)$ is, by definition, equal to $f(nx)$ for all x in $[0, 1]$. Since $V_0^1 g_n = V_0^n f$, (6) follows from Pólya's result.

4. The final result. We now state and prove a sharpened version of Theorem 3.

THEOREM 4. *If f is a real-valued function defined on $[0, \infty)$ such that* $\sup\{V_0^n f: n=1, 2, \cdots\} < \infty$, *then $\sum_{k=1}^\infty f(k)$ and $\int_0^\infty f(t)dt$ converge or diverge together.*

Proof. The plan of the proof is to express f in the form $f_1 - f_2$ where f_1 and f_2 are nonnegative, nonincreasing functions on $[0, \infty)$, and then to prove that the difference

$$\int_0^n f(t)\,dt - \sum_{k=1}^n f(k), \tag{7}$$

which we denote by d_n, converges as n approaches infinity.

To this end, let $P(x)$ and $N(x)$ denote the positive variation and negative variation, respectively, of f over the interval $[0, x]$, and denote $\sup\{V_0^n f: n=1, 2, \cdots\}$ by V. Then define f_1 and f_2 by the equations

$$f_1(x) = V + f(0) - N(x),$$
$$f_2(x) = V - P(x)$$

if $f(0) \geqq 0$ and by the equations

$$f_1(x) = V - N(x),$$
$$f_2(x) = V - f(0) - P(x)$$

if $f(0) < 0$. The reader can easily verify that $f = f_1 - f_2$ and that f_1 and f_2 are both nonnegative and nonincreasing on $[0, \infty)$.

Next, suppose that g is a nonnegative, nonincreasing function defined on $[0, \infty)$, and let

$$A_n = \int_0^n g(t)\,dt - \sum_{k=1}^n g(k) \qquad (n = 1, 2, \cdots).$$

Since $A_{n+1} - A_n = \int_n^{n+1} g(t)dt - g(n+1) \geqq 0$, $\{A_n\}$ is a nondecreasing sequence.

The obvious relation

$$\int_0^n g(t)\,dt = \sum_{k=0}^{n-1}\int_k^{k+1} g(t)\,dt \leqq \sum_{k=0}^{n-1} g(k)$$

implies that $A_n \leqq g(0)-g(n) \leqq g(0)$ $(n=1, 2, \cdots)$, and, consequently, the sequence $\{A_n\}$ converges.

Writing (7) in the form

$$\left\{\int_0^n f_1(t)\,dt - \sum_{k=1}^{n} f_1(k)\right\} - \left\{\int_0^n f_2(t)\,dt - \sum_{k=1}^{n} f_2(k)\right\}$$

and letting f_1 and f_2 play the rôle of g, we see that $\lim_{n\to\infty} d_n$ exists.

If $\int_0^\infty f(t)dt$ converges, then so does $\sum_{k=1}^\infty f(k)$ since

$$\sum_{k=1}^{n} f(k) = \int_0^n f(t)\,dt - d_n \qquad (n = 1, 2, \cdots).$$

Conversely, suppose that $\sum_{k=1}^\infty f(k)$ converges. Since $f=f_1-f_2$ where f_1 and f_2 are nonnegative, nonincreasing functions on $[0, \infty)$, $\lim_{t\to\infty} f(t)$ exists and is equal to 0. This last fact, in conjunction with the existence of $\lim_{n\to\infty}\int_0^n f(t)dt$, implies the convergence of $\int_0^\infty f(t)dt$, as desired.

5. Conclusion. We observe that (2), (4), and (6) give bounds on the absolute value of the difference (7) which hold even in the event of divergence. A knowledge of such bounds is often important for applications. For an example of a situation in which an extension of the integral test is employed, we refer the reader to a recent paper of D. J. Newman and H. S. Shapiro [**5**, pp. 251–252].

In conclusion, let us note that we have not stated results in their most general forms (e.g., for complex-valued functions) and that we have studiously ignored the question of integral tests for double series (cf. [**4**, p. 111]).

This work was supported by the National Science Foundation through Grant GP 1086.

References

1. Tom M. Apostol, Mathematical Analysis, Addison-Wesley, Reading, Mass., 1957.

2. Raymond W. Brink, A new sequence of integral tests for the convergence and divergence of infinite series, Ann. Math. 21 (1919) 39–60.

3. G. H. Hardy, Theorems connected with Maclaurin's test for the convergence of series, Proc. London Math. Soc., 9 (1910) 126–144.

4. G. H. Hardy, J. E. Littlewood, and G. Pólya, Inequalities, 2nd ed., Cambridge University Press, 1952.

5. D. J. Newman and Harold S. Shapiro, The Taylor coefficients of inner functions, Michigan Math. J., 9 (1962) 249–255.

6. G. Pólva and G. Szegö, Aufgaben und Lehrsätze aus der Analysis I, Springer, Berlin, 1925.

AN URN PROBLEM RELATED TO THE BALLOT PROBLEM

S. G. Mohanty, SUNY at Buffalo

The urn problem. Let an urn be given which contains n cards, marked with nonnegative integers $k_1, k_2, \cdots, k_n$ respectively where $k_1+k_2+\cdots+k_n=k$ and $0 \leqq k \leqq n$. Suppose that all the n cards are drawn without replacement from the urn. Denote by ν_j $(j=1, 2, \cdots, n)$ the number on the card drawn at the jth draw. What is the probability that

$$\nu_1+\nu_2+\cdots+\nu_r < r \quad \text{for } r = 1, 2, \cdots, n?$$

It is shown in [6] that

$$P\{\nu_1+\nu_2+\cdots+\nu_r < r \text{ for } r = 1, 2, \cdots, n\} = 1 - \frac{k}{n}, \tag{1}$$

a particular case of which is the classical "ballot problem." The same result has been derived by elegant combinatorial approaches in [1] and [3]. We shall obtain (1) by an elementary lattice path representation of the problem.

An interpretation. Let $N\{S\}$ denote the number of elements in S. Then

$$\begin{aligned} &N\left\{\sum_{i=1}^{r} \nu_i < r \text{ for } r = 1, 2, \cdots, n \,\middle|\, \sum_{i=1}^{n} \nu_i = k\right\} \\ &= N\left\{\sum_{i=1}^{r} \nu_i \leqq r-1 \text{ for } r = 1, 2, \cdots, n \,\middle|\, \sum_{i=1}^{n} \nu_i = k\right\} \\ &= N\left\{\sum_{i=1}^{r} (\nu_i+1) \leqq 2r-1 \text{ for } r = 1, 2, \cdots, n \,\middle|\, \sum_{i=1}^{n} (\nu_i+1) = k+n\right\} \\ &= N\left\{\sum_{i=1}^{r} \nu_i' \leqq 2r-1 \text{ for } r = 1, 2, \cdots, n \,\middle|\, \sum_{i=1}^{n} \nu_i' = k+n\right\}, \end{aligned} \tag{2}$$

where $\nu_i' = \nu_i+1$.

When $k=n$, the set is empty and hence we restrict our discussion to the case where $0 \leqq k < n$. Among nonnegative integers $k_1, k_2, \cdots, k_n$, let there be n_1 zeros, n_2 ones, and so on. Then $\sum i n_{i+1} = k$. It can be immediately seen from the definitions of compositions of an integer and the relation of domination defined on the compositions [5] that (2) is the number of n-compositions of m with given components of n_1 ones, n_2 twos, and so on, such that $\sum n_i = n$, $\sum i n_i = m$, where $n \leqq m \leqq 2n-1$, dominated by the n-composition

$$(1, \underbrace{2, 2, \cdots, 2}_{n-1}) \text{ of } 2n-1.$$

A lattice path representation. We provide below an alternative proof of (1) by using a correspondence between the urn problem and a set of lattice paths.

Let us represent the sequence $(\nu_1, \nu_2, \cdots, \nu_n)$ of numbers on the n cards drawn from the urn by a minimal lattice path in the following manner: (1) the path starts from the origin; (2) for every j, ν_j represents one horizontal unit followed by ν_j vertical units and the section of the path contributed by ν_j starts where the section of the path contributed by ν_{j-1} ended.

For instance, the path represented by (0, 0, 2, 1) is from $A(0, 0)$ to $P(3, 4)$ as shown in Figure 1. Note that the path from P to Q is a repetition of the path from A to P.

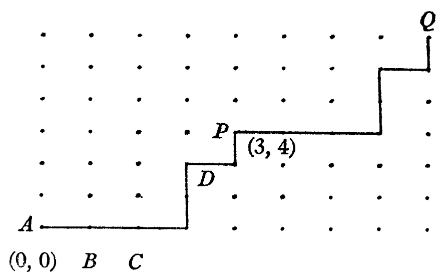


Fig. 1

It is obvious from the representation that the set $\{\nu_1+\nu_2+\cdots+\nu_r<r$ for $r=1, 2, \cdots, n\}$ is in 1:1 correspondence with the set of lattice paths from (0, 0) to (n, k) (where

$$k = \sum_{i=1}^{n} \nu_i \quad \text{and} \quad 0 \leqq k \leqq n)$$

with any order of $(\nu_1, \nu_2, \cdots, \nu_n)$ as vertical components such that no part of any path touches the line $x=y$ except at the origin.

We use a method similar to that of the penetrating analysis described in [4] for determining the required probability.

In the above example, any 7-unit segment path starting at either A or B or C or D represents a sequence of numbers which is a cyclic permutation of (0, 0, 2, 1). Think of the path APQ as an opaque screen and imagine an infinitely distant light source in the positive direction of the line $x=y$. Then A is the only lattice point among A, B, C, and D which would be in the light and therefore the path starting from A is the only one which would not touch the line $x=y$.

In general, the total number $n!$ of paths can be grouped into $(n-1)!$ classes, each consisting of n paths which are obtained by cyclic permutation of each other. Out of n points, which are the starting points of the n paths, obtained as cyclic permutations of each other, k points would be in the shade and consequently $n-k$ points would be in the light. If $k<n$, each of these points is the beginning of an $(n+k)$-unit segment path which will not touch the line $x=y$. Thus the probability is $(n-k)/n$. For $k=n$, the result holds trivially.

Referring to the definitions of a cyclic set of random variables and an ex-

changeable set of random variables in [2], we remark that the above method helps us to prove (1) essentially for a cyclic set of nonnegative integer-valued random variables $\nu_1, \nu_2, \cdots, \nu_n$. It is the same result as (1.3) in [2]. Since (1) is independent of the permutations of the random variables, it is therefore true for an exchangeable set of nonnegative integer-valued random variables $\nu_1, \nu_2, \cdots, \nu_n$.

COROLLARY. *If μ is a positive integer, and $0 \leqq \mu k \leqq n$, then*

$$P\left\{\nu_1 + \nu_2 + \cdots + \nu_r < \left[\frac{r-1}{\mu}\right] + 1 \text{ for } r = 1, 2, \cdots, n\right\} = 1 - \frac{\mu k}{n}.$$

The proof is obvious by noting that

$$P\left\{\nu_1 + \nu_2 + \cdots + \nu_r < \left[\frac{r-1}{\mu}\right] + \text{ for } r = 1, 2, \cdots, n\right\}$$

$$= P\left\{(\nu_1 + \nu_2 + \cdots + \nu_r) \leqq \left[\frac{r-1}{\mu}\right] \text{ for } r = 1, 2, \cdots, n\right\}$$

$$= P\{\mu(\nu_1 + \nu_2 + \cdots + \nu_r) \leqq r - 1 \text{ for } r = 1, 2, \cdots, n\}$$

$$= 1 - \frac{\mu k}{n}, \text{ since } \mu(\nu_1 + \nu_2 + \cdots + \nu_n) = \mu k.$$

References

1. M. Dwass, Fluctuation theory for integer-valued random variables, AFOSR Report No. 60–1036, (1960), Northwestern University.

2. ———, A fluctuation theorem for cyclic random variables, Ann. Math. Statistics, 33 (1962) 1450.

3. R. L. Graham, A combinatorial theorem for partial sums, Ann. Math. Statistics, 34 (1963) 1600.

4. G. D. Grossman, Fun with lattice points, Scripta Math., 16 (1950) 119.

5. T. V. Narayana, and G. E. Fulton, A note on the compositions of an integer, Canadian Math. Bull., I (1958) 169.

6. L. Takács, A generalization of the ballot problem and its application in the theory of queues, J. Amer. Stat. Assoc. 57 (1962) 327.

NEW GEOMETRIC REPRESENTATION OF AN OLD INFINITE SERIES

A. FEINGOLD, University of Missouri at Rolla

A familiar infinite series is

$$1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + - \cdots = \log 2. \tag{1}$$

From (1) follows

$$1/2 + 1/12 + 1/30 + \cdots = \log 2. \tag{2}$$

The area bounded by the curve $y = 1/x$, and the lines $x = 1$, $x = 2$, and $y = 0$

is therefore equal to the sum of the series

$$1/2 + 1/12 + 1/30 + \cdots.$$

The terms in this series can be represented by rectangles having, respectively, areas 1/2, 1/12, 1/30, etc., as shown in Fig. 1. The line $J'J$ is drawn first, creating the rectangle $WJ'JZ$ of area 1/2. Then AJ' is bisected at E', and the lines $E'E$ and ET are drawn to form the rectangle $J'E'ET$ of area 1/12. Now, two horizontal lines through G' and C' bisect $E'J$ and AE', respectively, leading to formation of rectangles $QGUT$ and $C'CNE'$ when GU and CN are drawn. The area of $QGUT$ is 1/30 and that of $C'CNE'$ is 1/56. The bisecting can be repeated indefinitely, always proceeding upwards in the construction of the rectangles, and the result is easy to foresee. Fig. 1 shows the next step and is self-explanatory.

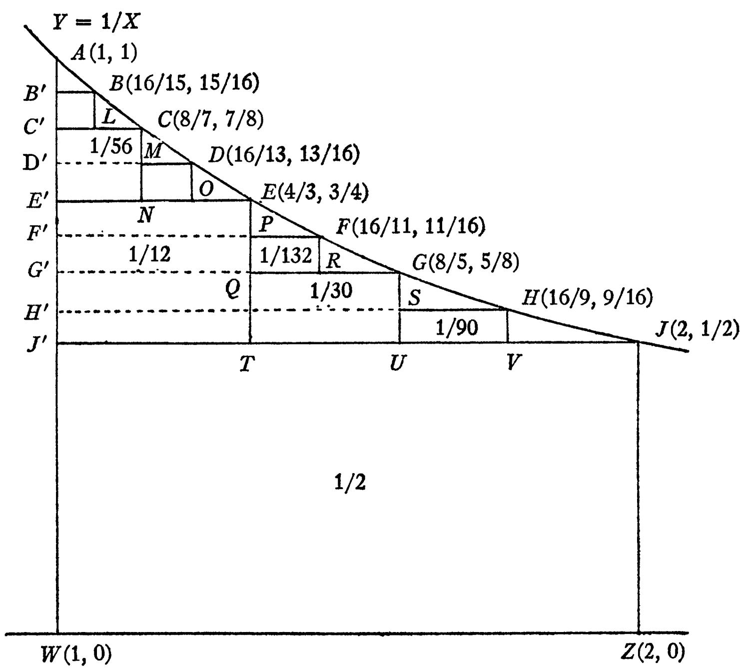


FIG. 1

Proof that the areas of rectangles obtained in the above manner continue to represent the successive terms of series (2) *indefinitely.*

Let us take a generic point on the curve $y=1/x$. Let its coordinates be p/q and q/p, p and q being mutually prime and p being equal to 2^n, while q is equal to 2^n-m, where m and n are integers and $m \leqq 2^{n-1}$. This generic point is now a

corner of one of the rectangles formed in the manner explained above. The other three corners have coordinates

$$(p/q, (q-1)/p),$$
$$(p/(q+1), (q-1)/p),$$
$$(p/(q+1), q/p),$$

respectively. The sides of this rectangle are $p/q - p/(q+1) = p/q(q+1)$ and $q/p-(q-1)/p = 1/p$. Thus the area is equal to $1/q(q+1)$. A little reflection shows that the denominator q of the x-coordinate of our generic point goes through all the odd numbers as the point itself is made to coincide in Fig. 1 first with J, then with E, G, C, H, F, D, etc. Therefore the largest rectangle has the area $1/1\cdot 2$, the next largest $1/3\cdot 4$, etc.

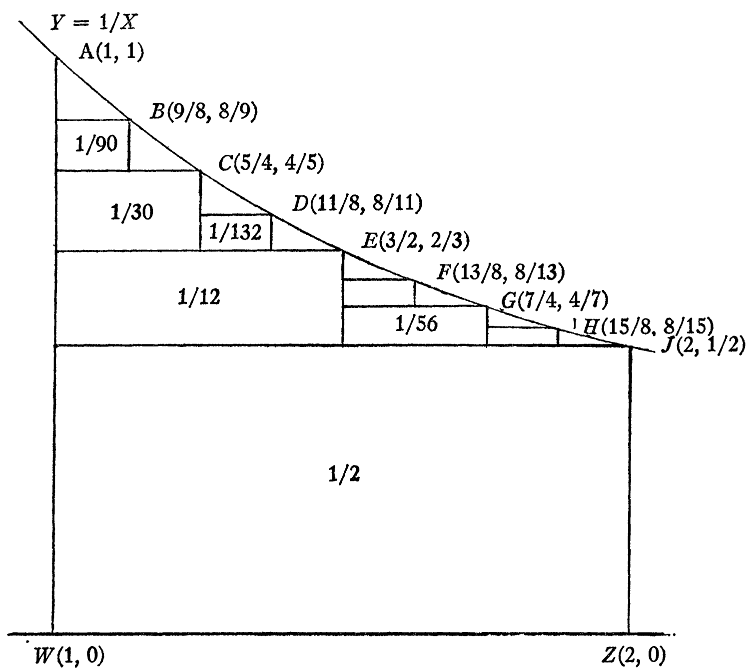


FIG. 2

A second representation of the same series is obtained using the subdivision shown in Fig. 2. The proof is similar to the one given above and is therefore omitted for brevity.

Another well-known sum is

$$\pi/4 = 1 - 1/3 + 1/5 - 1/7 + 1/9 - 1/11 + - \cdots,$$

and therefore

$$\pi/4 = 2/3 + 2/35 + 2/99 + \cdots.$$

The author has not found to date any geometric representation of the second series along the lines demonstrated above.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College, and
JOHN A. BROWN, University of Delaware

All material for this department should be sent to John R. Mayor, 1515 Massachusetts Avenue, N.W., Washington, D. C. 20005.

AN AFRICAN EDUCATION PROGRAM

HUGH P. BRADLEY, Educational Services, Inc., Newton, Massachusetts

In 1965 Educational Services Incorporated initiated two new activities under its African Education Program. In February, in Kano, Northern Nigeria, a conference of African, American and British educators during a five day discussion took the first steps in starting an African Elementary Science Program aimed at creating science units designed for use in African elementary classes. It was planned that the science teaching units should be suitable to the African economics and localities, and emphasize questioning and practical work and use of materials and "things" as opposed to the traditional rote learning from teachers and books. Later in the year, at Entebbe, Uganda, in July and August a similar multiracial group at a six-week Workshop produced—in various stages of teaching readiness, some twenty-two experimental units for trial teaching in African elementary schools.

To supervise this local trial teaching and to assist in encouraging local adaptations of the Entebbe science units, the new African Elementary Science Program is co-operating with African Ministries of Education and Universities in setting up Science Centers which will undertake local control of the reform of the science curriculum. So far two Science Innovators, Mr. William S. Warren, Jr., and Mr. Michael B. R. Savage, have been appointed to Science Centers in Nairobi and Nsukka respectively. Plans are well advanced for additional centers in Tanzania, Malawi and Mid-West Nigeria. It is anticipated that some ten or more of these centers will be set up in the next two years.

In the meantime plans are progressing for a second Workshop to be held—in Dar Es Salaam, Tanzania—in the summer of 1966.

In its second new venture, under its African Mathematics Program, E.S.I. is about to undertake the first main stage in making possible the full use of

modern mathematics in African elementary and secondary schools. The Ford Foundation has agreed to support a two-stage Institute which is aimed at making available to the countries participating in the Mathematics Program, a core of people knowledgable about modern mathematics, able to evaluate programs in modern mathematics, and capable of undertaking the training of other tutors, teachers and students in the content of modern mathematics courses and the activity methods involved in teaching them. Another aim of the institute is to make available materials which will enable this small group of key people to undertake the training of others at a lower level of mathematical ability. Each participating country is being invited to send to the Institute three senior mathematics tutors who will continue to be involved in teacher training work in the future, and one Ministry official likely to continue to be involved in the supervision of mathematics teaching in the schools. It is proposed to give these two sets of educators parallel courses beginning with a four-week course in July 1966, continuing by means of correspondence work and seminars during 1966 and 1967, and a second four-week course in July 1967, followed by more correspondence work into 1968.

Ten African mathematicians from African universities have been invited to assist in the Institute as seminar and discussion leaders and as controllers of the on-going correspondence work in 1966/67/68. They will also be doing some research work in their own mathematical specialty in conjunction with Professor A. Gleason of Harvard University, who will be the main lecturer. Professor D. E. Richmond of Williams College, Professor B. J. Pettis of the University of North Carolina, Chapel Hill, and Dr. Grace Alele Williams of Lagos University, Nigeria, will be the other lecturers at the Institute.

In the meantime the main program is continuing to produce experimental material for African elementary and secondary schools and for teacher training colleges. Last year at the Workshop held in Mombasa, Kenya, African and American educators and mathematicians wrote additional experimental textbooks for the series which is being tried in nearly eight hundred schools in the ten participating countries. By the end of 1965 experimental textbooks were available for the first four years of the elementary and the first four years of a five-year secondary course, and three volumes for teacher training colleges. Evaluation materials for the secondary and primary classes have also been prepared and are at various stages of pre-testing. In addition, in 1965 an experimental textbook for the first year of a four-year secondary course was produced. Finally, the Mathematics Program has been organizing teacher training institutes at which teachers are trained to use the experimental textbooks and supervisors are introduced to the "new" mathematics. Over the Christmas 1965 and New Year 1966 period, Institutes were held in Zambia, Tanzania (2), Nigeria and Liberia.

A changing feature of these institutes is that local personnel are now becoming available to assist in the lecturing, and in two cases it has not been necessary to send lecturers from the United States. If this trend can be accelerated by the

ABC Institute the program will be able to devote its full energies to the larger task of assisting the countries to prepare their tutors and colleges for the training of teachers for the full use of modern mathematics materials throughout the school systems.

SECONDARY SCHOOL MATHEMATICS CURRICULUM IMPROVEMENT STUDY

H. F. FEHR, Teachers College, Columbia University

Through a grant given by the U. S. Office of Education, Teachers College, Columbia University will inaugurate an experiment in curriculum construction in secondary school mathematics. A new detailed syllabus for a unified six-year program of mathematical education, grades seven through twelve, will be developed. The construction is to be free of any restrictions of traditional content or sequence. The syllabus is to encompass all the mathematics that is now considered essential through a first year university program, specifically the fundamental structures of number systems and of algebra, linear algebra, probability and its applications, mathematics related to computers, and the calculus (analysis).

From the syllabus complete textual material will be written by mathematicians and well-informed educators for use in a small number of pilot classes where the experimental teaching will be done by scholarly teachers. It is proposed not only to create a practical curriculum, but also to determine the teacher education in mathematics and in pedagogy necessary for the successful teaching of the program, and to develop tests and testing techniques to determine the extent and depth of mastery of the materials that are taught.

The experiment is under the directorship of Professor Howard F. Fehr, Teachers College, Columbia University. The major consultants for the program are Professor Marshall H. Stone, University of Chicago; Professor E. Ray Lorch, Columbia University; and Dr. Julius H. Hlavaty, National Council of Teachers of Mathematics. In June 1966, a group of eight outstanding mathematicians from the U.S.A., four distinguished European mathematicians, who have played a major role in creating modern, unified curriculums in their own countries, and four mathematical educators will meet to produce the proposed syllabus, based on a previously prepared position paper. During the summer of 1966, the textual material for grade seven will be written, and twenty-five teachers will be given 100 hours of preparatory training—50 hours in basic mathematical theory and 50 hours in the pedagogy—of the proposed materials. During the 1966–1967 school year, ten pilot classes (seven seventh grades and three eighth grades) will be taught, each class having two teachers. During this period the entire program will be subject to testing and analysis. At the end of the first year of teaching, a detailed report will be given to the mathematical profession.

It is the expressed hope to continue the experiment year by year throughout all high school grades.

ASSISTANCE FOR STUDENTS GIFTED IN MATHEMATICS

Teachers of high school students extraordinarily gifted in mathematics frequently arrange conferences for these students with a mathematician at a college or university in the vicinity. In communities where no college or university is easily accessible, teachers are invited to contact the School Mathematics Study Group Committee on the Extraordinarily Gifted Student for assistance in arranging individual study in mathematics for such students.

Address requests to: Committee on the Extraordinarily Gifted Student, School Mathematics Study Group, School of Education, Stanford University, Stanford, California.

BRIEF COMMENT

Report on In-service Training for Teachers of Mathematics, issued by the *Joint Mathematical Council of the United Kingdom*, London, June 1965.

The Joint Mathematical Council of the United Kingdom, a council of representatives from various British mathematical associations and societies, in an effort to improve the in-service training of school mathematics teachers, has issued this report. It is felt that although present efforts are laudable, they do not reach enough teachers, and an expansion of such a program on a more organized basis is necessary.

With the understanding that the number of teachers who require immediate help is large, that *all* teachers of mathematics require such training throughout their teaching careers, that time must be provided to allow the teacher to pursue such programs, that there is a shortage of adequately trained people to organise and staff such in-service training programs, and that there should be constant interchange of information between teachers and organisers of these programs, the Council has made the following recommendations:

"I. That as a matter of urgency facilities be provided within the school structure for serving teachers to develop their knowledge and efficiency, preferably by setting up Mathematics Centres under Local Education Authorities.

II. That each Institute of Education set up an Advisory Unit to assist all the agencies helping the in-service training of teachers of mathematics within its Region.

III. That a permanent sub-committee of the Schools Council for the Curriculum and Examinations be set up to work for the development of in-service training facilities for teachers of mathematics.

IV. That a National Information Centre be established to provide

(a) a communication link between all people concerned with the teaching of mathematics in schools;

(b) a source of information on all developments and publications in the field of mathematical education."

Copies of this report were distributed by The Mathematical Association, 22 Bloomsbury Square, London, W.C. 1, England, as one of the constituent societies of the council.

Emotional Perils of Mathematics, DONALD R. WEIDMAN, Letter to the Editor of *Science*, September 3, No. 3688, 149 (1965) 1048.

This is an unusually perceptive letter detailing some of the personal problems that mathematicians face when they devote themselves to mathematical research. The writer classifies the problems under four headings: the need for total involvement, frustration at obtaining no results, lack of appreciation, and self-dissatisfaction. He contrasts the problems that the research mathematician faces with problems that researchers in other fields face; and points out that the experimentalist has at least, as the results of his efforts, some data which have some meaning, whereas the mathematician may have nothing whatever after the expenditure of great effort. This letter ought to be required reading for graduate students who contemplate a career devoted to mathematical research.

Basic Library List. Committee on the Undergraduate Program in Mathematics of the Mathematical Association of America, 1965.

"This list of some 300 books, from which approximately 170 are to be chosen to form a basic library in undergraduate mathematics, is intended to do the following:

1. Provide the student with introductory material in various fields of mathematics which he may not have encountered previously.
2. Provide the student, whose interest has been aroused by his teachers, with reading material collateral to his course work.
3. Provide the student with reading at a level beyond that ordinarily encountered in his undergraduate curriculum.
4. Provide the faculty with reference material.
5. Provide the general reader with elementary material in the field of mathematics."

The list was produced over a period of several years through a combined effort, first by the Library Committee, and then by the Advisory Group on Communications of CUPM. It should be in the hands of every department chairman, and of such other faculty members as are interested in providing minimum library facilities in their field. Copies of the list are available at no cost from CUPM, P.O. Box 1024, Berkeley, California 94701.

The Entebbe Mathematics Workshop Summer 1962, ONYERISARA UKEJE, in A Report of an African Education Program, Educational Services Inc., 1965, 17–19.

The aim of the workshop, as it evolved, was "to make available to African countries several new discoveries and new changes that have taken place recently both in the content and the method of teaching school mathematics. The emphasis in method was to be on discovery." The workshop is an interesting and encouraging instance of cooperation among American professors and teachers of mathematics; British officials, professors, and teachers of mathematics; and African government officials, lecturers, and teachers of mathe-

matics. The production of the group encompasses material at both the primary and the secondary level and includes as well some objective testing material. "The Secondary and Primary materials are currently being tried out in several schools in West and East Africa with very encouraging reports thus far."

The Entebbe Mathematics Workshop Summer 1963, JOHN O. OYELESE, in A Report of an African Education Program, Educational Services Inc., 1965, 20–24.

The report details the efforts of a group of American, British, and African educators, officials and mathematicians to further the work of the 1962 Entebbe Workshop in producing materials for both primary and secondary grades, for testing, and for teacher training institutions. "For the first time an attempt was made to translate the Primary teachers' guides into Swahili, the language used over a large area of East Africa. The important lesson learned from the translation was that care should be exercised in writing the teachers' guides, and that only words which do not present any conceptual difficulties should be used."

"Africa has been rather fortunate to have the Entebbe Workshops. We are in the forefront of experiments going on in various parts of the world in the teaching of the 'New Mathematics.' In the past our ideas for the development of education came to us through Britain; and although developments in the teaching of mathematics are being made in Great Britain at the present time, it would normally have taken several years for these new ideas to reach the English-speaking countries in Africa. Now we do not need to lag behind; we can move forward together."

The Entebbe Mathematics Workshop Summer 1964, CYRIL N. OKOSI, in A Report of an African Education Program, Educational Services Inc., 1965, 25–29.

The report gives further information on the work begun by the 1962 and 1963 Entebbe Workshops in developing material for use in African schools both at the elementary and secondary levels, as well as testing materials and teacher training materials. "The aim of Entebbe project has been to make mathematics teaching and learning in African schools more interesting and more meaningful." Reports were also received on the use of the materials produced earlier in African schools. In general these reports seem quite favorable regarding the products of earlier workshops. For example, "it was hoped that in seven to ten years' time Entebbe mathematics would be used throughout Tanzania." Or again in regard to an institute held in January at Nsukka, Nigeria, "the institute was a huge success." Some difficulty seems to have developed in planning the secondary school geometry program, but agreement appears to have been reached on the direction this program should take. At this point three years of materials for primary grades, three years of materials for secondary grades, and

three volumes for teachers' training schools are available as the result of the Entebbe Workshops.

African Mathematics Program: Tutor and Teacher Training Institutes December, 1963 to February, 1965, STANLEY D. WEINSTEIN, in A Report of an African Education Program, Educational Services, Inc., 1965, 36–46.

This is "a brief report of fourteen courses conducted by Educational Services, Incorporated in tropical Africa between December, 1963, and February, 1965. The objective of these courses was to familiarize training college mathematics tutors and primary school teachers with the new approach to mathematics using as an example the textual materials written at the Entebbe Mathematics Workshops of ESI." Comments from both participants and lecturers are given at some length, and the general flavor is one of considerable enthusiasm; the institutes seem to have generated considerable pressure in several African countries toward marked alteration in the school mathematics curriculum and for the early introduction of the "new mathematics." Educational Services Incorporated is seeking funds in order to carry out a rather extensive program for further institutes of this kind in order to further re-education in mathematics in the shortest time possible in Africa.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE, Bloomfield College

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, Bloomfield College, Bloomfield, N. J. 07003. Proposers of Problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

All solutions of Elementary Problems should be sent to A. E. Livingston, University of Alberta, Edmonton, Canada. To facilitate their consideration, solutions of Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before September 30, 1966.

E 1885. *Proposed by D. E. Daykin and C. J. Eliezer, University of Malaya, Kuala Lumpur*

For which continuous functions $f(x)$ do we have

$$a_1^{f(a_1)} a_2^{f(a_2)} \cdots a_n^{f(a_n)} \geqq 1,$$

for all positive real numbers $a_1, \cdots, a_n$ with product $a_1a_2 \cdots a_n = 1$?

E 1886. *Proposed by W. M. Sanders, Lawrence University, Appleton, Wis.*

Show that if $ax^2+by^2+cz^2+2fxy+2gzx+2hyz+2rx+2sy+2tz+d=0$ represents two distinct planes, then the matrix

$$E = \begin{pmatrix} a & f & g & r \\ f & b & h & s \\ g & h & c & t \\ r & s & t & d \end{pmatrix}$$

is not the adjoint of any matrix.

E 1887. *Proposed by Simeon Reich, Haifa, Israel*

Let I, O, H, M, respectively, denote the incenter, circumcenter, orthocenter, and centroid of an acute-angled triangle ABC. Show that

$$9\frac{AM^2 + BM^2 + CM^2}{AO^2 + BO^2 + CO^2} - \frac{AH\cdot BH\cdot CH}{AI\cdot BI\cdot CI} \leqq 8,$$

with equality if and only if the triangle is equilateral.

E 1888. *Proposed by C. C. Lindner, Coker College, Hartsville, South Carolina*

Prove that if $n>2$, then $\sigma(n) \leqq n^{3/2}$, where $\sigma(n)$ is the sum of divisors of n.

E 1889. *Proposed by Joseph Arkin, Spring Valley, N. Y.*

Prove: The sum of 2^p squares $(\sum_{j=1}^{2^p} y_j^2)$ can always be expressed (in an infinite number of ways) as the quotient $(\sum_{j=1}^{2^p} A_j^2/\sum_{j=1}^{2^p} x_j^2)$ of two numbers, each of which is a sum of 2^p squares $(p=0, 1, 2, \cdots)$.

E 1890. *Proposed by Emanuel Vegh, U. S. Naval Research Laboratory.*

If m and n are given integers greater than one, show that there exist positive integers $N_1, N_2, \cdots, N_{m-1}$ such that

$$\sqrt{m} = 1 + \sum_{i=1}^{m-1} (\sqrt{N_i} - \sqrt{N_i - 1})^{1/n}.$$

E 1891. *Proposed by A. E. Livingston, University of Alberta*

Let f be a number-theoretic function, and set

$$F(n) = \sum_{d|n} f(d)f(n/d) \qquad (n = 1, 2, 3, \cdots).$$

If f is multiplicative (i.e., $f(mn) = f(m)f(n)$ when $(m, n) = 1$), it is easily verified that F is also multiplicative. Prove the conditioned converse: If $f(1) = 1$ and F is multiplicative, then f is multiplicative. (Compare with the well-known Theorem 6–8 in W. J. LeVeque, *Topics in Number Theory*, vol. 1, p. 88.)

E 1892. *Proposed by Melvin Hausner and Paul Magriel, New York University.*

An experiment has n possible outcomes σ_i, each with probability $p_i>0$, $\sum p_i=1$. The experiment is repeated for as many times (N) as necessary until all of the n different outcomes occur. What is the expected value of N?

E 1893. *J. H. Reid, Andover, England*

Imagine a square lattice with $a\times b$ vacant holes each one unit distant from its nearest neighbors. There exists an unlimited supply of n different species of atom, which repel each its own kind, and they are indistinguishable from one another for the purposes of this problem. How many arrangements exist so that no two atoms of the same species have to occupy adjacent holes?

E 1894. *Proposed by K. E. Whipple, Auburn University*

A student thought that the inverse of a function is the same as the reciprocal of the function. Are there any functions for which this is true? Considering only real valued functions that are continuous everywhere except zero, characterize those functions f for which $f^{-1}=1/f$.

SOLUTIONS OF ELEMENTARY PROBLEMS

E 1714 Returns

E 1766 [1965, 315]. *Proposed by Stanton Philipp, Long Beach, California*

For positive integers n and x define $N_n(x)$ to be the number of positive integers d such that (1) d divides x and (2) $x\leqq d^2\leqq n^2$. Prove that as $n\to\infty$, $n^{-2}\sum_{x=1}^{n^2} N_n(x)$ approaches $\frac{1}{2}$.

Solution by T. M. Apostol, California Institute of Technology. Let S denote the set of lattice points (x, d) satisfying the conditions $d\mid x$, $x\leqq d^2\leqq n^2$, $1\leqq x\leqq n^2$. Then we have

$$\sum_{x=1}^{n} N_n(x) = \sum_{(x,d)\in S} 1.$$

Since $d\mid x$, for each (x, d) in S we can write $x=qd$ with $1\leqq q\leqq d$, and we find

$$\sum_{(x,d)\in S} 1 = \sum_{d=1}^{n}\sum_{q=1}^{d} 1 = \sum_{d=1}^{n} d = \frac{n(n+1)}{2}\sim\frac{1}{2}\,n^2 \quad\text{as}\quad n\to\infty.$$

Note that if a weight $f(x/d)$ is attached to each lattice point (x, d) in S the same argument gives

$$\sum_{(x,d)\in S} f(x/d) = \sum_{d=1}^{n}\sum_{q=1}^{d} f(q) = \sum_{d=1}^{n} F(d),$$

where $F(d)=\sum_{q=1}^{d} f(q)$. Thus, for example, if $T_n(x)$ denotes the sum of those integers d satisfying conditions (1) and (2) then $f(q)=q$, $F(d)=d(d+1)/2$, and

$\sum_{d=1}^{n} F(d) = n^3/6 + n^2/2 + n/3$, whence $\sum_{x=1}^{n^2} T_n(x) \sim n^3/6$ as $n \to \infty$. More generally, if k is a nonnegative integer and if $T_{n,k}(x)$ denotes the sum of the kth powers of those d satisfying (1) and (2), we find

$$\sum_{x=1}^{n^2} T_{n,k}(x) \sim \frac{n^{k+2}}{(k+1)(k+2)} \quad \text{as} \quad n \to \infty.$$

Also solved by J. Beidler, G. Beumer (Netherlands), E. O. Buchman, G. E. Engebretsen, D. M. Hancasky, D. R. Hayes, Agnis Kaugars, R. L. Lipsman, Robert Maas, J. B. Muskat, W. M. Patterson & A. R. Wohlgemuth, Donald Quiring, R. L. Robinson, Robin Sibson (England), Guy Torchinelli, W. C. Waterhouse, R. C. Weger, and all the solvers of E 1714.

The editors regret the error in filing which permitted this item to appear a second time.

A Number-theoretic Sum

E 1767 [1965, 315]. *Proposed by H. W. Gould, West Virginia University*

If μ is the Möbius function and ϕ is Euler's totient function, show that $\sum_{d|n} \mu(d)\phi(d) = 0$ if and only if n is even.

Solution by E. S. Langford, U. S. Naval Postgraduate School. Let $F(n) = \sum_{d|n} \mu(d)\phi(d)$. Since μ and ϕ are multiplicative functions, F is likewise. (See Niven & Zuckerman, *An Introduction to the Theory of Numbers*, p. 84 ff. Cf. also problem 5, p. 89.) If p^m is the power of a prime, then $F(p^m) = 1 - (p-1) = 2 - p$. Therefore, in general, $F(n) = \prod_{p|n} (2-p)$, from which the result is evident.

Also solved by A. N. Aheart, T. M. Apostol, Y. M. ben-David, Ralph Bennett, P. M. Berry, M. A. Bershad, M. G. Beumer (Netherlands), E. O. Buchman, J. P. Burling, J. A. Burslem, R. M. Caron, Gary Chartrand, M. Chowdhury (Germany), D. I. A. Cohen, G. E. Engebretsen, R. B. Eggleton, P. K. Garlick, A. H. Gioia, D. M. Hancasky, D. R. Hayes, Agatha Himmelfarb & H. N. Wilson, Mrs. E. M. Horadam (Australia), Bernard Jacobson, J. A. Joseph, Geoffrey Kandall, M. S. Klamkin, C. C. Lindner, Andrzej Makowski (Poland), D. C. B. Marsh, M. G. Murdeshwar, Robert Patenaude, Harsh Pittie, R. L. Robinson, P. J. Ryan, P. A. Scheinok, Robin Sibson (England), R. Sivaramakrishnan (India), Sidney Spital, Robert A. Smith, Guy Torchinelli, A. M. Vaidya, C. Van de Vyle (Netherlands), W. C. Waterhouse, R. C. Weger, Leonard Weinstein, C. F. Wells, Barbara A. Welsh, B. D. Wick, D. S. Zave, David Zeitlin, and the proposer.

Editorial Note. Without appealing to multiplicative functions, the proof can easily be found as follows. $\phi(d)$ is even for all values of d except 1 and 2, which implies that each term of $\mu(d)\phi(d)$ is even for $d > 2$. $\mu(1)\phi(1) = 1$, $\mu(2)\phi(2) = -1$. Therefore $\sum \mu(d)\phi(d)$ is always odd except when terms for both $d = 1$ and $d = 2$ are included, i.e. when 2 is a divisor of n.

Conversely, if d is an odd divisor of an even integer n, then $\phi(2d) = \phi(d)$, $\mu(2d) = -\mu(d)$, $\mu(4d) = 0$; whence $\mu(2d)\phi(2d) + \mu(d)\phi(d) = 0$. Thus, neglecting divisors of n which are multiples of 4 (and contribute nothing to the sum) all other terms in the sum can be paired off so as to make the total sum $= 0$.

Identities for the Triangle

E 1768 [1965, 315]. *Proposed by Jean M. Quoniam, Saint-Etienne, France*

In the triangle ABC let D, E, F be the points of contact of the sides with the incircle of radius r. Let H be the orthocenter of the triangle DEF, s the area of

DEF, and I its incenter. Prove that

$$\sum \cos A = 3/2 - \overline{IH}^2/2r^2, \qquad \sum \sin A = 2s/r^2.$$

Solution by D. C. B. Marsh, Colorado School of Mines. In the statement, I should be the circumcenter of DEF (the incenter of ABC).

$\sum \cos A = -\sum \cos 2D$	(angles of ABC are supplements of twice the angles of DEF)
$= 4\prod \cos D + 1$	(familiar identity for angles of a triangle)
$= \dfrac{2r_0}{R_0} + 1$	(§299, Johnson, *Modern Geometry*: r_0 being the radius of the circle inscribed in the pedal triangle of DEF; $R_0 = r$ being the circumradius of DEF)
$= \dfrac{3}{2} - \dfrac{R_0^2 - 4r_0R_0}{2R_0^2}$	(algebraically identical to the preceding expression)
$= \dfrac{3}{2} - \dfrac{\overline{IH}^2}{2r^2}$	(§324, Johnson).
$\sum \sin A = \sum \sin 2D$	(relation of angles in ABC and DEF)
$= 4\prod \sin D$	(familiar identity for triangles)
$= \dfrac{\prod d}{2R_0^3}$	(circumdiameter equals ratio of side to sine of opposite angle)
$= 2s/r^2$	(product of sides equals 4 times product of area by circumradius; $R_0 = r$).

Also solved by G. E. Engebretsen, Mrs. A. C. Garstang, and D. M. Hancasky.

Coefficients of the Cyclotomic Polynomial

E 1769 [1965, 315]. *Proposed by I. J. Schoenberg, University of Pennsylvania*

Let $F_n(x)$ be the cyclotomic polynomial, i.e. the monic polynomial whose zeros are the primitive nth roots of unity. Show that all coefficients of $F_n(x)$ are nonnegative if and only if n is a power of a prime.

Solution by Sister Marion Beiter, Rosary Hill College, Buffalo, N. Y.

If n is prime, $F_n(x) = (x^n - 1)/(x-1) = \sum_{j=0}^{n-1} x^j$.

If $n = p^t$, p prime, $F_n(x) = F_p(x^{p^{t-1}})$.

Thus, if n is a power of a prime, the coefficients of $F_n(x)$ are exclusively 0's and 1's.

Now let $n = p_1^{t_1} p_2^{t_2} \cdots p_r^{t_r}$, $r > 1$. $F_n(x)$ has the term $x^{\phi(n)}$, because it is of degree $\phi(n)$, and has the constant term 1, the product of the primitive nth roots of unity. Furthermore, $F_n(1) = 1$ when n has at least two distinct prime factors (See Nagel: *Introduction to Number Theory*). Thus $F_n(x)$ has two coefficients

equal to 1 while the sum of all coefficients is $F_n(1)=1$, implying that $F_n(x)$ has a negative coefficient.

Also solved by D. M. Bloom, L. Carlitz, D. I. A. Cohen, V. H. Keiser, Jr., and the proposer.

Radially Symmetric Region

E 1770 [1965, 315]. *Proposed by G. P. Graham and T. A. Porsching, Carnegie Institute of Technology*

Let C be a planar simple closed curve enclosing a region R which is starlike in point P. Then a necessary and sufficient condition that C be radially symmetric in P is that every line through P divides R into subregions of equal area. Also, every line dividing R into equal areas passes through P.

Solution by Werner Baron, Vienna, Austria. We introduce a polar coordinate system; the pole is P, the polar axis is an arbitrary fixed half-line issuing from P. The periodicity of ϕ implies $r(\phi)=r(\phi+2\pi)$. The planar simple closed curve C is starlike in P, therefore we have for every ϕ one and only one value $r(\phi)$.

(i) For every α let

$$\frac{1}{2}\int_{\alpha}^{\alpha+\pi} r^2(\phi)\cdot d\phi - \frac{1}{2}\int_{\alpha+\pi}^{\alpha+2\pi} r^2(\phi)\cdot d\phi = 0. \tag{1}$$

This means that every line through P divides R into subregions of equal area. The derivative of (1) with respect to α must be zero for all α, too, because of the validity of (1) for all α.

$$\tfrac{1}{2}(r^2(\alpha-\pi) - r^2(\alpha) - r^2(\alpha+2\pi) + r^2(\alpha+\pi)) = 0,$$
$$r(\alpha+\pi) = r(\alpha).$$

This means that C is radially symmetric in P.

(ii) Let C be radially symmetric in P, i.e. $r(\phi+\pi)=r(\phi)$. Substituting $\phi=\psi+\pi$ in $\frac{1}{2}\int_{\alpha+\pi}^{\alpha+2\pi} r^2(\phi)\cdot d\phi$ we get

$$\frac{1}{2}\int_{\alpha+\pi}^{\alpha+2\pi} r^2(\phi)\cdot d\phi = \frac{1}{2}\int_{\alpha}^{\alpha+\pi} r^2(\psi)\cdot d\psi$$

for all α. This means every line through P divides R into two equal areas, completing the proof.

For the second part of the problem we give an indirect proof. Let g be a line which divides R into two subregions A and B with equal area but which does not pass through P. Let A be the subregion which contains P. There exists a line h parallel to g which passes through P. Let D be the subregion of R between h and g. (We can use the same letters for the subregions and for their areas.) P is an inner point of R and so the part of R between h and g must be of positive area. Therefore $D>0$, $A=B$ and $A-D=B+D$, giving an evident contradiction.

Also solved by E. O. Buchman, D. I. A. Cohen, Robert Connelly, G. E. Engebretsen, Neal Felsinger, D. M. Hancasky, E. S. Langford, K. W. Reed, Jr., and Robin Sibson (England).

Sums of Subsets of a Set of n Integers

E 1771 [1965, 316]. *Proposed by Andy Vince, Stanford University*

Given a set of n positive integers, $a_1, a_2, \cdots, a_n$, it is always possible to choose a subset of these integers such that the sum of its elements is divisible by n.

Solution by P. K. Garlick, New Mexico State University. All congruences are modulo n. Let

$$s_i = a_1 + a_2 + \cdots + a_i \qquad i = 1, 2, \cdots, n.$$

If $s_i \equiv s_j$ for some $i > j$, then $s_i - s_j = a_{j+1} + \cdots + a_i \equiv 0$. If not, then the s_i's form a complete residue system and there exists $s_k = a_1 + \cdots + a_k \equiv 0$.

Also solved by T. M. Apostol, Robert Bart, Y. M. ben-David, Ralph Bennett, M. A. Bershad, Walter Bluger, D. A. Breault, J. L. Brown, Jr., E. O. Buchman, T. J. Burke, J. W. Cannon, Gary Chartrand, M. R. Chowdhury (Germany), Allan Chuck & Peter Goldstein, D. I. A. Cohen, Robert Connelly, J. C. Deel, L. E. De Noya, J. F. Dillon, G. E. Engebretsen, A. L. Epstein, N. J. Fine, J. H. Foster & K. M. Hunter, Wallace Growney, R. B. Hardin, Jr., R. H. Hines, Jr., W. Imrich (Austria), Bernard Jacobson, R. A. Jacobson, James Joseph, Geoffrey Kandall, J. H. B. Kemperman, Charles McCracken, E. L. Magnuson, Andrzej Makowski (Poland), D. C. B. Marsh, Weston Meyer, J. E. Motzkin, J. B. Muskat, Sam Newman, R. J. Oberg, F. D. Parker, B. J. Parshall, Jagdish Prasad, Donald Quiring, Azriel Rosenfeld, Albert Schild, Robin Sibson (England), D. L. Silverman, Richard Sinkhorn, M. W. Sterling, M. V. Tamhankar (India), W. C. Waterhouse, R. C. Weger, J. C. Williams, and the proposer.

Several solvers note that the restriction of positiveness is not necessary for the a_i's. Moreover, as shown in the proof, the selected subset can be required to have consecutive indices.

D. C. B. Marsh offers the following more general phraseology for the problem: Let $a_1, a_2 \cdots a_n$ be the "product" of any n elements, not necessarily all distinct, of a finite group of order n, then some subproduct $a_j a_{j+1} \cdots a_k$ $(i \leqq j \leqq k \leqq n)$ is the group identity. This, as Makowski observes, is the substance of problem 4300 [1950, 47].

Co-Invertibility of Some Square Matrices

E 1772 [1965, 316]. *Proposed by Henry Cox, David Taylor Model Basin, Washington, D. C.*

Let A and B be arbitrary matrices of dimensions $m \times n$ and $n \times m$ respectively. Does the existence of $[I_m + AB]^{-1}$ imply the existence of $[I_n + BA]^{-1}$, where I_m and I_n are the identity matrices of dimension m and n respectively? If so, determine the latter in terms of the former. If not, give a counterexample.

I. *Solution by A. S. Householder, Oak Ridge National Laboratory.* The determinants of the following three matrices are all equal, hence all are singular or all are nonsingular:

$$\begin{pmatrix} I_n & A \\ B & I_m \end{pmatrix}, \qquad I_m + AB, \qquad I_n + BA.$$

II. *Solution by K. L. Yocom, South Dakota State University.* More generally if k is a nonzero scalar then $AB - kI_n$ has an inverse if and only if $BA - kI_n$

has an inverse. To prove this, suppose $\det(AB-kI_m)=0$. Then k is an eigenvalue of AB and there is a nonzero m-dimensional column vector x such that $ABx=kx$. Multiplying by B on the left, $BA(Bx)=kBx$. Now $Bx\neq 0$, for otherwise $A(Bx)=A(0)=0=kx$, a contradiction. Thus Bx is an eigenvector of BA corresponding to the eigenvalue k, and $\det(BA-kI_n)=0$. A similar argument establishes the converse.

Let $C=AB-kI_m$ and $D=BA-kI_n$ be nonsingular. Then $BC=DB$ and $D^{-1}B=BC^{-1}$. Multiplying on the right by A and adding $-kD^{-1}$, it follows that $D^{-1}BA-kD^{-1}=BC^{-1}A-kD^{-1}$ or $D^{-1}D=BC^{-1}A-kD^{-1}=I_n$. Hence

$$D^{-1} = \frac{1}{k}\, BC^{-1}A - \frac{1}{k}\, I_n.$$

Setting $k=-1$, the desired result is obtained.

Also solved by S. Baron, Joel Brawley, Jr., David Carlson, Yu Chang, D. I. A. Cohen, Peter Enis, N. J. Fine, D. R. Hayes, L. N. Howard, P. W. M. John, Geoffrey Kandall, P. G. Kirmser, Murray Lieb, D. C. B. Marsh, Dave Nixon, I. Olkin, Robin Sibson (England), M. V. Tamhankar (India), D. H. Underwood, W. C. Waterhouse, R. C. Weger, Alan Weinstein, Lenard Weinstein, R. E. Wheeler, T. A. Whitelaw (England), and J. Ernest Wilkins, Jr.

Lenard Weinstein bases one solution on the following theorem on p. 105 of N. Jacobson' *Lectures in Abstract Algebra:* If A is an $m\times n$ matrix and B is an $n\times m$ matrix, then the characteristic polynomial of BA is x^{n-m} times the characteristic polynomial of AB.

John and Wheeler note that the result may be found in K. D. Tocher, *The design and analysis of block experiments*, J. Royal Stat. Soc., Series B, 14 (1952) 46.

Triangles Whose Sides Are Consecutive Integers

E 1773 [1965, 316]. *Proposed by Michael Lieber, AVCO Corporation and Harvard University*

Let T be a triangle whose sides are consecutive integers and whose area is an integer. Prove that one of the altitudes divides the triangle into two Pythagorean right triangles. Furthermore, this altitude divides the base into segments whose lengths differ by 4. (The 1-2-3 "triangle" and the 3-4-5 right triangle are the only degenerate cases.)

Solution by M. V. Tamhankar and M. B. Suryanarayana, India. Let $n-1$, n, $n+1$ be the lengths of the sides AB, BC, CA of T. The area is $\Delta=\frac{1}{4}n\sqrt{(3n^2-12)}$ by Hero's formula. Now n is even, otherwise Δ is not integral. Say $n=2m$, whence $\sqrt{(3m^2-3)}$ is the (integral) length of the altitude AD.

Also $\overline{DC}^2=(2m+1)^2-3(m^2-1)=(m+2)^2$, giving $DC=m+2$ and $BD=m-2$. Finally, $DC-BD=4$ and, since all segments are integral, the triangles are Pythagorean.

Also solved by A. N. Aheart, Nancy C. Baggs & Dorothy J. Stodola, Robert Bart, M. A. Bershad, M. G. Beumer (Netherlands), D. A. Blaeuer, W. Bluger, J. A. Burslem, L. Carlitz, Richard Chandler, Allan Chuck, Kay S. Clever, D. I. A. Cohen, Monte Durham, R. B. Eggleton (Australia), G. E. Engebretsen, Neal Felsinger, S. M. Gagola, Jr., Michael Goldberg, S. H. Greene, Louise S. Grinstein, Wallace Growney, Stephen Hoffman, Agnis Kaugars, Sam Kravitz, E. S.

Langford, Esther A. Linfield, D. C. B. Marsh, Gus Mavrigian, Norman Miller, Robert Patenaude, C. B. A. Peck, Harsh Pittie, K. W. Reed, Jr. & Nelda Viser & J. C. Jones, Simeon Reich, P. A. Scheinok, D. R. Shoemaker, Robin Sibson (England), Al Somayajulu, E. M. Stone, Adrian Struyk, Guy Torchinelli, Simon Vatriquant (Belgium), Van de Vyle (Belgium), M. C. Weinrich, Lenard Weinstein, Hazel S. Wilson, R. I. Winston & J. B. Small, K. L. Yocom, and David Zeitlin.

Editorial Note. The problem has been given before. See Dickson, *History of the Theory of Numbers*, v. II, pp. 191–202, for many references to this and related problems. Note in particular, problem 4047 [1944, 102–104] where formulas are derived giving all values of n. In fact, if n_j represents the jth value in the sequence of all values of n, we will have

$$n_j = (2+\sqrt{3})^j + (2-\sqrt{3})^j, \qquad n_{j+2} = 4n_{j+1} - n_j.$$

The first few values of n (besides 2 and 4) are 14, 52, 194, 724.

Monotonicity of log $(ax+1)/\log(bx+1)$

E 1774 [1965, 316]. *Proposed by Wayne E. Smith, University of California, Los Angeles*

Let $0<a<b$. Prove that $\log(ax+1)/\log(bx+1)$ is a monotonic increasing function for all $x>0$.

I. *Solution by P. M. Berry, Socony Mobil Oil Co., Dallas, Texas.* After converting the derivative of the given function to a proper fraction, it is seen that the denominator is always positive and the numerator is

$$g(x) = a(bx+1)\log(bx+1) - b(ax+1)\log(ax+1).$$

Now $g'(x)=ab\log[(bx+1)/(ax+1)]>0$. Therefore $g(x)$ is strictly monotonic increasing on $(0, \infty)$. As $g(0)=0$, it follows that $g(x)>0$ for $x>0$. Therefore the derivative of the given function is always positive so that $\log(ax+1)/\log(bx+1)$ is strictly monotonic increasing.

II. *Solution by Ingram Olkin, Stanford University.* We note the following more general framework, assuming the existence of the relevant derivatives.

If $g(x)=f(a,x)/f(b,x)$, $a<b$, then $g(x)$ is monotonic increasing if and only if

$$h(a,x) \equiv \frac{\partial \log f(a,x)}{\partial x} > \frac{\partial \log f(b,x)}{\partial x} \equiv h(b,x),$$

which holds if and only if $\partial h/\partial a<0$. Thus $g(x)$ is increasing if and only if $(\partial^2/\partial a\partial x)\log f(a,x)<0$.

For the present case $f(a,x)=\log(ax+1)$, $0<a$, $0<x$,

$$\frac{\partial^2 \log f(a,x)}{\partial a\partial x} = \frac{\log(ax+1) - ax}{(ax+1)^2[\log(ax+1)]^2}$$

which is negative because $\log(1+y)<y$, $y>0$.

Also solved by Y. M. ben-David, J. L. Brown, Jr., T. J. Burke, F. A. Butter, Jr., M. R. Chowdhury (Germany), D. I. A. Cohen, W. O. Egerland, G. E. Engebretsen, A. D. Fine, K. R. Goodearl, S. H. Greene, Harry Guess, D. R. Hayes, G. A. Heuer, Stephen Hoffman, R. A. Jacobson, James Joseph, M. S. Klamkin, E. S. Langford, H. C. Lauer, Murray Lieb, D. C. B. Marsh,

Norman Miller, Dave Nixon, Robert Patenaude, J. R. Porter, B. E. Rhoades, L. A. Ringenberg, P. A. Scheinok, Robin Sibson (England), Al Somayajulu, G. P. Speck, Sidney Spital, M. V. Tamhankar & M. B. Suryanarayana (India), Simon Vatriquant (Belgium), W. C. Waterhouse, Lenard Weinstein, Barbara A. Welsh, and the proposer.

Observations with Zero Dispersion

E 1775 [1965, 316]. *Proposed by George Purdy, University of Reading, England.*

Under what conditions do real $x_1, \cdots, x_n$ satisfy the equation $x_1^2 + \cdots + x_n^2 = (x_1 + \cdots + x_n)^2/n$ for $n \geqq 1$?

I. *Solution by J. M. Perry, Wells College, Aurora, N. Y.* Denoting by $\bar{x}$ the mean of the numbers x_i, the result follows immediately from the standard statistical identity

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 \equiv \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2.$$

Thus we have $\sum x_i^2 = (\sum x_i)^2/n$ only when $\sum (x_i - \bar{x})^2 = 0$, which requires $x_i = \bar{x}$ for all i; i.e., all of the numbers x_i must be equal.

II. *Solution by F. A. Butter, Jr., California State College, Long Beach.* If each $x_k > 0$, and $G(t) \equiv [n^{-1}(x_1^t + \cdots + x_n^t)]^{1/t}$, the given equation takes the form $nG^2(2) = nG^2(1)$, or $G(1) = G(2)$. It is known that $0 < G(a) \leqq G(b)$ if $0 < a < b$, and $G(a) = G(b)$ only if x_i are constant.

III. *Solution by M. S. Klamkin, Ford Scientific Laboratory, Dearborn, Mich.* The problem here is a special case of a well-known result for convex functions, i.e., if $\phi(t)$ is convex in $t \geqq 0$, then

$$\phi\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \leqq \frac{\phi(x_1) + \phi(x_2) + \cdots + \phi(x_n)}{n},$$

$(x_i \geqq 0)$, with equality only if the x_i's are equal or $\phi(t)$ is linear. For the present case $\phi(t) = t^2$ and we must have $x_1 = x_2 = \cdots = x_n$.

Also solved by A. N. Aheart, T. M. Apostol, R. A. Avelsgaard, J. H. Avila, Werner Baron (Austria), Robert Bart, M. A. Bershad, M. G. Beumer (Netherlands), W. J. Blundon, Joel Brawley, Jr., J. L. Brown, Jr., E. O. Buchman, Allan Chuck, D. I. A. Cohen, D. M. Cohen, Kenneth Dieter, G. C. Dodds, W. O. Egerland, G. E. Engebretsen, Peter Enis, A. L. Epstein, P. K. Garlick, Michael Goldberg, S. H. Greene, D. M. Hancasky, Carl Harris, D. R. Hayes, G. A. Heuer, Stephen Hoffman, J. C. Hohman, Bernard Jacobson, R. A. Jacobson, James E. Joseph, Erwin Just, S. C. King, Richard Kowalski, E. S. Langford, Murray Lieb, B. V. Limaye, Charles McCracken, D. M. Mahamunulu, A. Makowski (Poland), D. C. B. Marsh, J. T. Mathis, Norman Miller, D. E. Moxness, R. J. Newman, Sam Newman, David Nixon & Dick George, Robert Patenaude, Harsh Pittie, J. R. Porter, Simeon Reich, J. P. Ruebsamen, Y. P. Sabharwal, (India), P. A. Scheinok, Robin Sibson (England), Richard Sinkhorn, David A. Smith, Mitchell Snyder, Al Somayajulu, Sidney Spital, R. E. Stockton, R. L. Syverson, M. V. Tamhankar & M. B. Suryanarayana (India), Guy Torchinelli, Simon Vatriquant (Belgium), Julius Vogel, W. C. Waterhouse, R. J. Weber, Lenard Weinstein, J. M. Wild, K. L. Yocom, D. E. Zave, David Zeitlin, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before November 30, 1966.

5390. *Proposed by D. W. Hight, Kansas State College of Pittsburg*

Is there a series $\sum_{i=1}^{\infty} u_i(x)$ of functions continuous on a closed interval $[a, b]$ that converges absolutely and uniformly for which the Weierstrass M-test fails?

5391. *Proposed by J. H. Halton and M. Skibinsky, Brookhaven National Laboratory*

Give an explicit characterization of the convex hull of the arc from $t=0$ to $t=1$ of the twisted cubic $x=t$, $y=t^2$, $z=t^3$.

5392. *Proposed by Hwa Hahn, Pennsylvania State University*

Let $f(x)$ be the number obtained from a real x by reading the representation of x in the base p as the representation in the base $q(1<p\leqq q)$. Evaluate $\int_0^1 f(x)dx$.

5393. *Proposed by A. A. Mullin, University of California, Livermore*

Does there exist an integer for which the $n\times n$ matrix whose elements are $a_{ij}=(i,j)$ is singular? (i,j) is the greatest common divisor of the natural numbers i and j.

5394. *Proposed by P. J. O'Hara, Jr. University of Miami, Coral Gables, Florida*

Show that the functions, $(1-z)/(1-z^n)$, $n=1, 2, \cdots$, are uniformly bounded in modulus on the disk $\{z: |z-\frac{1}{2}|\leqq\frac{1}{2}\}$ in the complex plane; and that in fact for all n the bound is attained at $z=0$.

5395. *Proposed by Roy Westwick and W. A. McWorter, University of British Columbia*

Suppose G is a two-person game satisfying:

(1) The players play alternately.

(2) There is no draw.

(3) There are at least one and at most N $(<\infty)$ plays a player can make.

Prove that either there is a strategy insuring a win for the first player, or there is a strategy insuring a win for the second player.

5396. *Proposed by Albert Wilansky, Lehigh University*

Find a set X and two normal (Hausdorff) topologies T and T', for X, such that every T compact set is T' closed, but not every T' compact set is T closed. (The proposer does not know whether this can be done with metrizable topologies.)

5397. *Proposed by R. E. Shafer, University of California, Livermore*

Show, for $|\mathrm{Re}(\nu)| < \frac{1}{2}$ and $|\arg z| < \pi$, that

$$K_\nu(z) = e^{-z} z^\nu \cos \pi\nu \int_0^\infty \frac{e^{-t}}{t+z} t^{-\nu} I_\nu(t)\, dt,$$

where I_ν, K_ν are Bessel functions.

5398. *Proposed by G. A. Heuer, Concordia College, Moorhead, Minnesota*

Let $\sum a_n$ be a conditionally convergent series, S the set of its partial sums, and S' the derived set (set of accumulation points) of S. As is well known, rearranging the series $\sum a_n$ may alter S'. Consider all possible rearrangements of the fixed series $\sum a_n$. What are the sets S' which occur?

5399. *Proposed by L. F. Shampine, Sandia Corporation, Albuquerque, N.M.*

Let $p(n)$ be the largest spectral radius attained by any (0, 1) matrix with n nonzero entries. Prove $p \sim \sqrt{n}$.

SOLUTIONS OF ADVANCED PROBLEMS

Classifying and Summing $(d, n/d)$

5290 [1965, 555]. *Proposed by Daniel I. A. Cohen, Princeton University*

Let $\tau(a)$ be the number of divisors of a, and let $\lambda(a)$ be the number of distinct prime divisors of a. Prove that

$$(1) \qquad \tau(n) = \sum_{k^2|n} 2^{\lambda(n/k^2)},$$

$$(2) \qquad \sum_{d|n} (d, n/d) = \sum_{k^2|n} k 2^{\lambda(n/k^2)}.$$

I. *Solution by C. S. Venkataraman, Sree Kerala Varma College, Trichur, India.* If r is the total number of distinct prime divisors of n, then the number of ways of expressing n as the product of two coprime factors is 2^{r-1}. Let $d|n$, $(d, n/d) = k$, $d = kd_1$, $n/d = kd_2$. Then $(d_1, d_2) = 1$, $n = k^2 d_1 d_2$, and $k^2|n$.

Conversely, let $k^2|n$, $n = k^2 m$. Then m can be factored into $m_1 m_2$ with $(m_1, m_2) = 1$ in 2^{s-1} ways, where s is the total number of distinct prime divisors of m. For each of these ways, $n = k^2 m_1 m_2 = (km_1)(km_2) = (d)(n/d)$ with $d = km_1$ and $n/d = km_2$.

It follows that, for each k satisfying $k^2|n$, there exist 2^{s-1} and only 2^{s-1} pairs of divisors d, n/d such that $(d, n/d) = k$. Thus, the total number of such divisors is $2 \cdot 2^{s-1} = 2^s = 2^{\lambda(n/k^2)}$.

The $\tau(n)$ divisors of n can now be grouped into a finite number of mutually exclusive classes $C_1, C_2, \cdots$ such that C_k contains the divisors d for which $(d, n/d) = k$, if $k^2|n$; then the number of divisors in C_k is plainly $2^{\lambda(n/k^2)}$. [Note that C_k is the empty class if $k^2 \nmid n$.] It follows that $\tau(n) = \sum_{k^2|n} 2^{\lambda(n/k^2)}$.

To prove (2) we use again the fact that there are $2^{\lambda(n/k^2)}$ divisors d for which

$(d, n/d)=k$, yielding at once

$$\sum_{d|n}(d, n/d) = \sum_{k^2|n} k\cdot 2^{\lambda(n/k^2)}.$$

II. *Solution by E. S. Langford, U. S. Naval Postgraduate School.* Define the arithmetical function f by $f(n)=2^{\lambda(n)}$. It is not hard to see that f is multiplicative, and hence that F is multiplicative, where $F(n)=\sum_{k^2|n} f(n/k^2)$. If p is any prime, $F(p^m)=m+1=\tau(p^m)$ and (1) now follows from the multiplicative nature of $\tau(n)$.

Both sides of (2) are also multiplicative. Write $G(n)=\sum_{d|n}(d, n/d)$ and suppose that $(m, n)=1$. Then if $d|mn$, d can be written as $d=d_1d_2$, where $d_1|m$ and $d_2|n$, implying $(d, mn/d)=(d_1d_2, (m/d_1)(n/d_2))=(d_1, m/d_1)(d_2, n/d_2)$. Therefore

$$G(mn) = \sum_{d|mn}(d, mn/d) = \sum_{d_1|m}\sum_{d_2|n}(d_1, m/d_1)(d_2, n/d_2) = G(m)\cdot G(n).$$

Likewise we can establish that $H(n)=\sum_{k^2|n} k2^{\lambda(n/k^2)}$ is multiplicative.

If p is a prime, then

$$G(p^m) = \sum_{k=0}^{m}(p^k, p^{m-k})$$

$$= \begin{cases} 2+2p+2p^2+\cdots+2p^{(m/2)-1}+p^{(m/2)} & (m \text{ even}) \\ 2+2p+2p^2+\cdots+2p^{(m-1)/2} & (m \text{ odd}). \end{cases}$$

$$H(p^m) = \begin{cases} 1\cdot f(p^m)+p\cdot f(p^{m-2})+\cdots+p^{(m/2)}\cdot f(1) & (m \text{ even}) \\ 1\cdot f(p^m)+p\cdot f(p^{m-2})+\cdots+p^{(m-1)/2}\cdot f(p) & (m \text{ odd}). \end{cases}$$

Since $f(p^k)=2$ when $k\geqq 1$, and $f(1)=1$, it follows that $G(n)=H(n)$ whenever $n=p^m$, and the proof of (2) is complete.

Also solved by B. N. Bajaj, Robert Breusch, E. O. Buchman, E. P. Emerson, A. E. Fekete, A. S. Fraenkel (Israel), Michael Greening (Australia), D. A. Hejhal, Edward Hook, C. C. Lindner, A. E. Livingston, Stanton Philipp, Harsh Pittie, Simeon Reich, Michael Rosen, S. K. Sehgal & V. C. Dumir, R. Sivaramakrishnan (India), E. W. Trost (Switzerland), D. A. Zave, and the proposer.

Rosen observes that formula (1) may be found in Kronecker, *Vorlesungen über Zahlentheorie.* Fekete and Trost call attention to Liouville, *Sur quelques fonctions numeriques*, J. de Math. (2), 2 (1857), p. 141 ff., where (1) appears—possibly for the first time, although Fekete observes that the problem may be originally due to Dirichlet who treated it in his 1840 *Zahlentheorie.* Other similar interesting formulas may be found in the Liouville paper. The methods of attack in these earlier works involved the multiplication techniques of Dirichlet series and generating functions—methods used by Emerson and Sivaramakrishnan in their solutions of 5290.

Variations of the Sum $\sum z^n/n^n$

5291 [1965, 555]. *Proposed by D. S. Mitrinović, University of Belgrade, Yugoslavia*

Establish the formula

$$\sum_{n=0}^{\infty} \frac{1}{(n+1)^{n+1}} = \int_0^1 x^{-x}\, dx,$$

and generalize to $\sum z^n/(n+1)^{n+1}$.

I. *Solution by Harsh Pittie, Swarthmore College.* Noting that

$$\int_0^1 (zx)^m (\log x)^n\, dx = \frac{-n}{m+1} \int_0^1 (zx)^m (\log x)^{n-1}\, dx,$$

we have a simple recursion: $\int_0^1 (zx)^n (\log x)^n dx = (-1)^n z^n n!/(n+1)^{n+1}$. Hence

$$\int_0^1 x^{-zx}\, dx = \int_0^1 e^{-zx \log x}\, dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^1 (zx \log x)^n\, dx$$
$$= \sum_{n=0}^{\infty} \frac{z^n}{(n+1)^{n+1}},$$

where the interchange of summation and integration is easily justified for all complex z.

II. *Solution by Louis Comtet, Viroflay, France.* Put

$$F_{\alpha,\beta}(z) = F(z) = \int_0^1 \exp\left\{ zx^\alpha \left(\log \frac{1}{x}\right)^\beta \right\} dx,$$

where $\alpha > 0$, $\beta > 0$, z (complex) are given. Clearly, $F(z) = \int_0^1 \exp\{z \cdot \phi(x)\} dx$, where $\phi(x) = x^\alpha (\log 1/x)^\beta$, if $x > 0$, and $\phi(0) = 0$. Further, since $|\phi(x)| \leqq (\beta/\alpha e)^\beta$, for $0 \leqq x \leqq 1$, we can write

$$F(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \left\{ \int_0^1 \phi^n(x) \cdot dx \right\}.$$

Now

$$\int_0^1 \phi^n(x)\, dx = \int_0^1 x^{\alpha n}(-\log x)^{\beta n}\, dx = \frac{1}{(\alpha n+1)^{\beta n+1}} \int_0^\infty e^{-t} \cdot t^{\beta n}\, dt = \frac{\Gamma(\beta n+1)}{(\alpha n+1)^{\beta n+1}}.$$

Consequently:

$$F(z) = \int_0^1 \exp\left\{ zx^\alpha \left(\log \frac{1}{x}\right)^\beta \right\} dx = \sum_{n=0}^{\infty} \frac{\Gamma(\beta n+1)}{(\alpha n+1)^{\beta n+1}} \cdot \frac{z^n}{n!}.$$

Putting $\alpha=\beta=1$, we obtain

$$F_{1,1}(z) = \int_0^1 x^{-xz}\,dx = \sum_{n=0}^{\infty} \frac{z^n}{(n+1)^{n+1}}\,.$$

Also solved by Om. P. Anand, Daniel Asimov, G. Baron & W. Imrich (Austria), Y. M. ben-David, Robert Breusch, E. O. Buchman B. J. Cerimele, J. A. Fancher, A. E. Fekete, W. D. Fryer, Edward Garelis, Mrs. A. C. Garstang, Rudolf Gorenflo (Germany), J. H. Halton, Eldon Hansen, Stephen Hoffman, A. S. B. Holland, R. A. Horn, James Kiefer, M. S. Klamkin, Ben Klein, A. G. Konheim, E. S. Langford, A. E. Livingston, J. R. McGregor, D. F. Mackie, D. P. Mather, Stanton Philipp, S. U. Rangarajan (England), H. J. Ricardo, Charles Ryavec, M. N. Sastry & B. V. Rao (India), Michael Schulz, Richard Sinkhorn, Sidney Spital, F. W. Steutel (Netherlands), Arun Verma (India), Benjamin Volk, L. E. Ward, Sr., and D. A. Zave,

Notes. Rangarajan cites another formula which generalizes the question of the problem:

$$\int_0^1 x^{-zx}x^{\beta}\left(\log\frac{1}{x}\right)^{\gamma} dx = \sum_{n=0}^{\infty} \frac{\Gamma(1+\gamma+n)}{n!}\,\frac{z^n}{(n+\beta+1)^{n+\gamma+1}}\,.$$

Our contributors find many occurrences of the stated formula. See Pólya & Szegö, *Aufgaben*, v. 1, no. 160; Edwards, *Integral Calculus*, p. 135; Jolley, *Summation of Series*; Bierens de Haan, *Tables of Definite Integrals*; Ryshik & Gradstein, *Tables;* Douglas Aircraft Co. Report SM-14642; and the London B.Sc., B.A. Special Examination Papers of 1956.

The Taylor Expansion of Arc sec x

5292 [1965, 555]. *Proposed by Eldon Hansen, Lockheed Missiles and Space Co., Palo Alto, California*

Obtain the Taylor series expansion for arcsec x about a point $x=a$.

I. *Solution by A. E. Fekete, Memorial University of Newfoundland.* This is problem no. 11 on page 84 in Edwards, *An Elementary Treatise on the Differential Calculus* (London, 1921), where three terms of the expansion are given as

$$\text{Arcsec}\, x = \text{Arcsec}\, a + \frac{1}{a\sqrt{(a^2-1)}}(x-a) - \frac{2a^2-1}{a^2(a^2-1)^{3/2}}\,\frac{(x-a)^2}{2!} + \cdots.$$

Clearly, the general term takes the form $f^{(n)}(a)(x-a)^{n+1}/(n+1)!$ with $f(x) = d$ Arcsec $x/dx=(x^4-x^2)^{-1/2}$, $|x|>1$.

The problem thereby is reduced to finding $f^{(n)}(a)$. This is possible by using the formula of Faá de Bruno:

$$\frac{d^nz}{dx^n} = \sum \frac{n!}{1!^{i_1}i_1!2!^{i_2}i_2!\cdots n!^{i_n}i_n!}\,\frac{d^{i_1+\cdots+i_n}z}{dy^{i_1+\cdots+i_n}}\left(\frac{dy}{dx}\right)^{i_1}\left(\frac{d^2y}{dx^2}\right)^{i_2}\cdots\left(\frac{d^ny}{dx^n}\right)^{i_n}$$

where the summation is to be extended to all partitions $1i_1+2i_2+\cdots+ni_n=n$ with the convention $0^0=1$.

In our case $z=y^{-1/2}$, $y=x^4-x^2$, and

$$\frac{d^iz}{dy^i} = i!\binom{-\frac{1}{2}}{i}y^{-1/2-i} = (-1)^i\,\frac{(2i)!}{2^{2i}i!}\,\frac{1}{x^{2i+1}(x^2-1)^{1/2+i}}\,,$$

$$dy/dx = 2x(2x^2 - 1), \qquad d^2y/dx^2 = 2!(6x^2 - 1),$$
$$d^3y/dx^3 = 3!(2^2x), \qquad d^4y/dx^4 = 4!; \qquad d^ky/dx^k = 0 \quad \text{for } k \geqq 5.$$

Thus only those partitions will give nontrivial contributions which have no summand greater than 4, i.e., $1i_1+2i_2+3i_3+4i_4=n$. Substituting into Faá de Bruno's formula, we get

$$f^{(n)}(a) = \sum_{1i_1+2i_2+3i_3+4i_4=n} (-1)^{i_1+i_2+i_3+i_4} \cdot \frac{(2i_1+2i_2+2i_3+2i_4)!}{i_1!i_2!i_3!i_4!(i_1+i_2+i_3+i_4)!2^{i_1+2i_2+2i_4}} \cdot \frac{(2a^2-1)^{i_1}(6a^2-1)^{i_2}}{a^{i_1+2i_2+i_3+2i_4+1}(a^2-1)^{1/2+i_1+i_2+i_3+i_4}}, \qquad |a| > 1.$$

II. *Solution by F. W. Steutel, Enschede, Netherlands.* Changing the notation slightly we have

$$\text{(1)} \quad \frac{d}{du}\operatorname{arcsec}(a+u) = \frac{1}{a+u}\{(a+u)^2 - 1\}^{-1/2} = \frac{1}{a+u}\int_0^\infty e^{-(a+u)t}I_0(t)\,dt.$$

(See e.g., Watson, *Theory of Bessel Functions*, 2nd ed., pp. 77 and 384.) Expansion of $(a+u)^{-1}e^{-(a+u)t}$ in powers of u yields

$$\frac{d}{du}\operatorname{arcsec}(a+u) = \sum_{n=0}^{\infty}(-1)^n u^n \sum_{k=0}^{n} \frac{a^{k-n-1}}{k!}\int_0^\infty e^{-at}t^k I_0(t)\,dt,$$

and therefore

$$\text{(2)} \quad \operatorname{arcsec}(a+u) - \operatorname{arcsec} a = \sum_{n=0}^{\infty}(-1)^n \frac{1}{n+1}\left(\frac{u}{a}\right)^{n+1} \sum_{k=0}^{n} \frac{a^k}{k!}\int_0^\infty e^{-at}t^k I_0(t)\,dt,$$

which may be written in terms of hyper-geometric functions (cf., Watson, p. 385):

$$\operatorname{arcsec}(a+u) - \operatorname{arcsec} a = \frac{1}{a}\sum_{n=0}^{\infty}\frac{1}{n+1}\left(\frac{u}{a}\right)^{n+1}\sum_{k=0}^{n} {}_2F_1\left(\frac{k+1}{2}, \frac{k+2}{2}, 1, \frac{1}{a^2}\right).$$

An alternative form of (2) (see (1)) is

$$\operatorname{arcsec}(a+u) - \operatorname{arcsec} a = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{u^n}{n!}\int_0^\infty e^{-at}t^{n-1}\bar{I}_0(t)\,dt,$$

where $\bar{I}_0(t) = \int_0^t I_0(x)dx$.

Also solved by J. M. Quoniam (France), and by the proposer.

Solving $\sum f(d)f(n/d)=1$

5293 [1965, 555]. *Proposed by Martin J. Cohen, Beverly Hills, California*

Find a function f such that $\sum f(d)f(n/d)=1$ for every positive integer n, where the sum is taken over all d which divide n (including 1 and n).

I. *Solution by Richard Stanley, California Institute of Technology.* More generally, consider $F(n)=\sum_{d|n} f(d)f(n/d)$, where $F(n)$ is a given multiplicative function, i.e., $F(mn)=F(m)F(n)$ when $(m, n)=1$. Then $F(n)$ is determined by its value at p^k, p prime. Of course, $F(1)=1$ always. Let $F_p(k)=F(p^k)$. If $f(1)=1$, then it may be proved that $f(n)$ is also multiplicative, so that we need to solve

$$F(p^k) = F_p(k) = \sum_{d|p^k} f(d)f(p^k/d) = \sum_{i=0}^{n} f_p(i)f_p(k-i)$$

for each unknown function f_p. (The only other solution will be $-f(n)$ corresponding to $f(1)=-1$.)

Let $G_{f_p}=\sum_{i=0}^{\infty} f_p(i)x^i$. Then, formally,

$$G_{f_p}^2(x) = \sum_{i=0}^{\infty} f_p(i)x^i = G_{F_p}(x).$$

In particular, if $G_{F_p}(x)$ represents a function analytic in some neighborhood of the origin, then the same is true of $G_{f_p}(x)$ (since $G_{F_p}(0)=1$). In this case we have

$$f_p(k) = k!\frac{d^k}{dx^k}(G_{F_p}(x))^{1/2}\bigg|_{x=0}.$$

Specializing to $F_p(k)=1$ for all p yields

$$G_{F_p}(x) = \sum_{i=0}^{\infty} x^i = (1-x)^{-1}, \qquad G_{f_p}(x) = (1-x)^{-1/2}$$
$$= \sum_{i=0}^{\infty}(-1)^i\binom{-\frac{1}{2}}{i}x^i, \qquad f_p(k) = (-1)^k\binom{-\frac{1}{2}}{k}.$$

Thus the solution to the stated problem is $\pm f$ where:

$$f(p_1^{a_1}\cdots p_r^{a_r}) = (-1)^{a_1+\cdots+a_r}\binom{-\frac{1}{2}}{a_1}\cdots\binom{-\frac{1}{2}}{a_r}.$$

Other multiplicative functions $F(n)$ which lead to simple values of $f(n)$ are
(i) $F(n)=\mu(n)$, $G_{F_p}(x)=1-x$, $f_p(k)=(-1)^k\binom{\frac{1}{2}}{k}$;
(ii) $F(n)=|\mu(n)|$, $G_{F_p}(x)=1+x$, $f_p(k)=\binom{\frac{1}{2}}{k}$;
(iii) $F_p(k)=1/k!$, $G_{F_p}(x)=e^x$, $f_p(k)=1/(2^kk!)$;
(iv) $F(n)=\tau(n)$, $G_{F_p}(x)=(1-x)^{-2}$, $f_p(k)=1$.
Note also that if $F(n)=\sum_{d|n} f(d)f(n/d)$, then

$$F(n)n^s = \sum_{d|n} f(d)d^s f(n/d)(n/d)^s,$$

so that, for example, if $F(n) = \mu(n)n^s$, then $f_p(k) = (-1)^k\binom{\frac{1}{2}}{k}p^{ks}$.

II. *Solution by M. S. Klamkin, Ford Motor Company.* The problem may be extended to find a function F such that

$$\sum_{d_1d_2\cdots d_r=n} F(d_1)F(d_2)\cdots F(d_r) = 1$$

for every positive integer n where the sum is taken over all d_r which divide n (including 1 and n).

Consider the formal product of r identical Dirichlet series:

$$\left\{\sum \frac{F(n)}{n^s}\right\}^r = \sum \frac{G(n)}{n^s}.$$

[See Hardy and Wright, Theory of Numbers, p. 248.] Then,

$$G(n) = \sum_{d_1d_2\cdots d_r=n} F(d_1)F(d_2)\cdots F(d_r) = 1,$$

whence,

$$\sum \frac{F(n)}{n^s} = \zeta(s)^{1/r} = \prod_{\text{primes}}\{1 - 1/p_n^s\}^{-1/r}$$

multiplied possibly by an rth root of unity. Let

$$\{1 - 1/p_n^s\}^{-1/r} = 1 + \frac{a_1}{p_n^s} + \frac{a_2}{p_n^{2s}} + \frac{a_3}{p_n^{3s}} + \cdots,$$

then $a_m = \binom{-1/r}{m}$. It now follows that if

$$n = p_{\alpha_1}^{i_1}\cdot p_{\alpha_2}^{i_2}\cdot\cdots\cdot p_{\alpha_s}^{i_s}, \quad \text{then} \quad F(n) = a_{i_1}a_{i_2}\cdots a_{i_s}$$

multiplied by a fixed rth root of unity. The original problem corresponds to the special case $r=2$.

Also solved by G. Baron & W. Imrich (Austria), W. J. Blundon, Robert Breusch, E. P. Emerson, R. C. Entringer, A. E. Fekete, N. J. Fine, W. D. Fryer, Michael Greening (Australia), R. P. Kelisky, James Kiefer, Donald Knuth, E. S. Langford, A. E. Livingston, O. P. Lossers (Netherlands), Stanton Philipp, S. K. Sehgal & V. C. Dumir, R. Sivaramakrishnan (India), M. W. Sterling, M. V. Subbarao, R. E. Walde, D. A. Zave, and the proposer.

Notes: (1) By using the form

$$4^{-\Sigma a_j}\prod\binom{2a_j}{a_j}$$

for the value of $f(n)$ (see I above), Lossers gives the formula $f(n) = (1/2)^{t_1}(3/4)^{t_2}(5/6)^{t_3}\cdots$, where t_i is the number of exponents a_j which are $\geq i$ in the prime representation of n.

(2) Fekete notes that for the function $f(n)$ in I we have the following relations:

$$F(n) = \sum_{d|n} f(d) = \binom{a_1 + \frac{1}{2}}{a_1}\binom{a_2 + \frac{1}{2}}{a_2}\cdots\binom{a_r + \frac{1}{2}}{a_r},$$
$$\tau(n) = \sum_{d|n} f(d)F(n/d).$$

The functions f, F, τ are special cases of

$$\tau_\alpha(n) = \binom{a_1 + \alpha}{a_1}\binom{a_2 + \alpha}{a_2}\cdots\binom{a_r + \alpha}{a_r}$$

with $\alpha = -\frac{1}{2}, \frac{1}{2}, 1$ respectively, which may be found treated in L. Carlitz, *Extended Stirling and Exponential Numbers*, Duke Math. J., 32 (1965) 205–224.

(3) Subbarao develops a theory of submultiplicative functions ($f(mn)-f(m)f(n)$ is of constant sign, or zero, m and n coprime) from which the solution of the problem follows.

Simultaneous Inequalities for Analytic Functions

5294 [1965, 555]. *Proposed by Frank Dean, Long Beach, California*

Let f and g be functions analytic at the point p. Show that the conditions:

$$|f'(p)|^2 > |f(p)f''(p)|, \qquad |g'(p)|^2 > |g(p)g''(p)|$$

guarantee the existence of a point q such that:

$$|f(q)| > |f(p)| \quad \text{and} \quad |g(q)| > |g(p)|.$$

Solution by James Kiefer, National Institutes of Health, Bethesda, Md. We shall show that for all θ in a closed interval $[\alpha, \alpha+\pi]$ there is a $\rho_0 = \rho_0(\theta) > 0$ such that for all $\rho < \rho_0$, $|f(p+\rho e^{i\theta})| > |f(p)|$. A similar situation obtains for $g(z)$. Then the two closed intervals must have a common element modulo 2π and it follows that for all q on some line segment starting at p we have $|f(q)| > |f(p)|$, $|g(q)| > |g(p)|$.

In the following, the function g may replace the function f. We note first that the hypothesis assures that $f'(q) \neq 0$ (and $g'(q) \neq 0$) in a neighborhood of p. If $f(p) = 0$, then this is an isolated zero and we have a neighborhood of p in which $|f(q)| > 0$. Thus, to establish the assertion of the first paragraph we need only now consider the case $f(p) \neq 0$.

We write $w(z) = f(p+zf(p)/f'(p))/f(p) = 1+z+\frac{1}{2}cz^2\phi(z)$ where $c = f(p)f''(p)/[f'(p)]^2$, $|c| < 1$, $\phi(0) = 1$, $\phi(z)$ analytic in a neighborhood of $z = 0$. Letting $z = te^{i\theta}$, we find $dw/dt = e^{i\theta}$ at $t = 0$ and it follows that for θ in the open interval $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, with $w = 1+te^{i\theta}+\frac{1}{2}ct^2e^{2i\theta}\phi(te^{i\theta})$, we have $|w| > 1$ for sufficiently small positive t. It remains to be shown that this is also true at the end points of the interval.

With $z = it$, $w = u+iv = 1+it-\frac{1}{2}ct^2\phi(t)$ we obtain, at $t = 0$:

$$\frac{dw}{dt} = i, \quad \frac{d^2w}{dt^2} = -c \equiv -\alpha - i\rho, \quad \frac{dv}{dt} = i, \quad \frac{du}{dt} = 0, \quad \frac{du}{dv} = 0,$$

and from $d^2u/dt^2 = (d^2u/dv^2)(dv/dt)^2 + (du/dv)(d^2v/dt^2)$ we see, further, that

$$-\alpha = \frac{d^2u}{dt^2} = -\frac{d^2u}{dv^2} \quad \text{or} \quad \left|\frac{d^2u}{dv^2}\right| < 1.$$

Thus in the case of the function $w(z)$ we see that a vertical segment of the z-plane through the origin is mapped into an arc in the w-plane passing through $w=1$, with vertical tangent and with curvature less than 1. Thus the arc in the w-plane is locally exterior to the unit circle except for the point of tangency. The proof is now complete.

Also solved by Robert Breusch, and by M. D. Mavinkurve (India).

Note. Mavinkurve shows that the conclusion follows for all q in a set with positive two-dimensional measure.

Factors of the Number of Divisors of $(f(n))!$

5295 [1965, 555]. *Proposed by M. V. Subbarao, University of Alberta, Canada*

Let a and k be arbitrary positive integers, and b, c positive integers such that $3b \geqq 5c$. Let $N=(kn^b(n^c+1))!$. If $t(N)$ denotes the number of divisors of N, show that for all sufficiently large positive integers n, $t(N)\equiv 0 \pmod{(1+k)^a}$.

Solution by Robert Breusch, Amherst College. More generally: Let a and k be arbitrary positive integers. Let $f(n)$ be *any* strictly increasing number-theoretic function. Let $N=(f(n))!$. Then, for all large enough n, $t(N)\equiv 0 \pmod{(1+k)^a}$.

Proof. A prime p with $p>\sqrt{f(n)}$ will be contained $[f(n)/p]$ times in N; thus it will contribute to $t(N)$ the factor $([f(n)/p]+1)$. Now $[f(n)/p]=k$ if and only if $k\leqq f(n)/p<k+1$, that is, iff $f(n)/(k+1)<p\leqq f(n)/k$. The prime number theorem implies that for every sufficiently large n, there will be at least a such primes, and again for n large enough, all these primes will be greater than $\sqrt{f(n)}$. Thus $t(N)$ will contain a factor $(1+k)^a$.

Also solved by the proposer.

Fitting the Weierstrass Polynomial

5296 [1965, 555]. *Proposed by J. E. Wetzel, University of Illinois*

The theorem of Weierstrass asserts that any continuous function f on an interval $[a, b]$ can be uniformly approximated with arbitrary accuracy by a polynomial.

(a) Let $x_1, \cdots, x_n$ be distinct points of $[a, b]$. Show that the polynomial p can be chosen to pass through all the points $(x_1, f(x_1)), \cdots, (x_n, f(x_n))$ in addition to approximating f uniformly on $[a, b]$.

(b) Let $x_1, \cdots, x_n$ be distinct points not in $[a, b]$ and $y_1, \cdots, y_n$ arbitrary numbers. Show that the polynomial p can be chosen to pass through all the points $(x_1, y_1), \cdots, (x_n, y_n)$ in addition to approximating f uniformly on $[a, b]$.

I. *Solution by L. J. Wallen, Stevens Institute of Technology, Hoboken, N. J.* (a) Let p be a polynomial with $p(x_i)=f(x_i)$, $i=1, \cdots, n$, and write $f-p=f^+-f^-$ where f^+ and f^- are the positive and negative parts of $f-p$. (Note:

$f^+(x_i)=f^-(x_i)=0$.) Pick polynomials q_ν and r_ν satisfying $q_\nu\to(f^+)^{1/n}$ and $r_\nu \to(f^-)^{1/n}$ uniformly. Setting

$$Q_\nu = \prod_{i=1}^{n} [q_\nu - q_\nu(x_i)] \quad\text{and}\quad R_\nu = \prod_{i=1}^{n} [r_\nu - r_\nu(x_i)],$$

we have $Q_\nu\to f^+$, $R_\nu\to f^-$. Finally $p+Q_\nu-R_\nu\to f$ and $p(x_i)+Q_\nu(x_i)-R_\nu(x_i)=f(x_i)$.

(b) This follows from (a) by extending f to an interval including the x_i so as to satisfy $f(x_i)=y_i$.

II. *Solution by Benjamin Volk, Yeshiva University.* Choose $Q(x)$ so that $|Q(x)-f(x)|<\epsilon$ for $a\leqq x\leqq b$. With $\lambda_k=f(x_k)-Q(x_k)$, $1\leqq k\leqq n$, let

$$R(x) = \sum_{k=1}^{n} \lambda_k \prod_{\substack{i=1\\ i\neq k}}^{n} \frac{x - x_i}{x_k - x_i}, \quad\text{and}\quad P(x) = Q(x) + R(x).$$

Then

$$|R(x)| \leqq \epsilon \sup_{a<x<b} \sum_{k=1}^{n} \prod_{\substack{i=1\\ i\neq k}}^{n} \left|\frac{x - x_i}{x_k - x_i}\right| = A\epsilon,$$

$$\begin{aligned}|P(x) - f(x)| &= |Q(x) + R(x) - f(x)| \leqq |Q(x) - f(x)| + |R(x)| \\ &\leqq (1 + A)\epsilon,\end{aligned}$$

which implies $P(x_k)=f(x_k)$ and this completes the proof.

Also solved by G. Baron & W. Imrich (Austria), Béla Bollobás (Hungary), Robert Bowen, Jon Cole, Michael Golomb, Harry Guess, James Kiefer, A. G. Konheim, M. D. Mavinkurve (India) L. D. Meeker, E. A. Nordgren, V. L. N. Sarma (India), John Shaw, G. P. Speck, F. W. Steutel (Netherlands), and the proposer.

Bowen, Meeker and Nordgren show that the result of the problem is an immediate consequence of an extension of the Stone-Weierstrass theorem: *The uniform closure of the set of polynomials vanishing on a fixed finite set in* $[a, b]$ *includes all the continuous functions of* $[a, b]$ *which vanish on the given set.* See M. H. Stone, A Generalized Weierstrass Approximation Theorem, p. 46 in *Studies in Modern Analysis*, vol. I, of the MAA.

Two Sided Ideals of a Regular Ring

5297 [1965, 556]. *Proposed by S. Lajos, K. Marx University of Economics, Budapest, Hungary*

Prove that any two-sided ideal J of a two-sided ideal I of a (von Neumann) regular ring R is again a two-sided ideal of R.

Solution by E. P. Emerson, University of Minnesota. I is itself a regular ring, because if $a\in I$, then $axa=a$ for some $x\in R$; now if $y=xax$, then $y\in I$ and $aya=ax(axa)=a$. Let $a\in J$ and $r\in R$. It follows that $ra\in I$ and there exists $x\in I$ such that $raxra=ra$. But $raxr\in I$ so $ra\in J$. The same argument shows that $ar\in J$.

Also solved by Robert Bowen, K. S. Eldridge, Harry Gonshor, Mary Gray, Erwin Just, Geoffry Kandall, Irving Katz, T. P. Kezlan, Kwangil Koh, W. G. Leavitt's class in Ring Theory, C. C. Lindner, Jiang Luh, R. K. Markanda (India), M. D. Mavinkurve (India), Barbara L. Osofsky, Harsh Pittie, J. R. Porter, A. Radhakrishna (India), Klaus Schmitt, E. A. Schreiner, Robert Shanny, R. P. Sullivan, and the proposer.

Professor Leavitt's class offers the following generalization: If J is an ideal of an ideal I of a ring R such that J is von Neumann regular in R, then J is an ideal of R.

When a Semigroup is a Group

5298 [1965, 556]. *Proposed by S. Lajos, K. Marx University of Economics, Budapest, Hungary*

A subsemigroup A is said to be an (m, n)-ideal of the semigroup S if $A^mSA^n \subseteq A$, where m, n are nonnegative integers and A^m is suppressed if $m=0$. Prove that a semigroup S is a group if and only if it has no proper (m, n)-ideal, where m, n are fixed positive integers.

Solution by E. A. Schreiner, Western Michigan University. Let m, n be fixed positive integers. Assume that S is a group, A an (m, n)-ideal of S, $a \in A$. Then $a^m(a^{-(m+n)})a^n = e \in A$. Thus for arbitrary $s \in S$, $e^mse^n = s \in A$, and so $A = S$.

Conversely, assume that S has no proper (m, n)-ideals. For arbitrary $a \in S$, aSa is a subsemigroup and since $(aSa)^mS(aSa)^n \subseteq aSa$, aSa is an (m, n)-ideal. Therefore $aSa = S$. Hence, for arbitrary $a, b \in S$, the equations $ax = b$ and $ya = b$ have solutions $x, y \in S$. This is sufficient to insure that S is a group.

Also solved by Mary K. Bennett, D. R. Brown, Michael Greening (Australia), U. S. Kahlon (India), Irving Katz, Kwangil Koh, C. C. Lindner, A. E. Livingston, M. D. Mavinkurve (India), Barbara L. Osofsky, J. R. Porter, A. Radhakrishna (India), K. N. Sigmon, R. P. Sullivan, and the proposer.

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University
COLLABORATING EDITORS: K. O. MAY, University of California, Berkeley, and E. P. VANCE, Oberlin College

Materials intended for review should be sent directly as follows: Books: R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457. Programmed Materials: K. O. May, University of California, Berkeley, Calif. 94704. Films: E. P. Vance, Oberlin College, Oberlin, Ohio 44074.

Computer Programming: A Mixed Language Approach. By Marvin L. Stein and William D. Munro. Academic Press, New York, 1965. 459 pp. $11.50.

In their preface, the authors state that the objective of their book is to provide training for both the professional programmer and the user interested in the computer only as a device to help solve mathematical problems. To this

end, they discuss the machine languages, symbolic machine languages, and problem-oriented languages.

Their exposition is clear, with several examples and exercises to assist the reader. A listing of the chapters will indicate the subject matter covered by the book: 1. Number Systems; 2. Machine Organization; 3. Elementary Coding; 4. Fixed and Floating Point Arithmetic: Scaling; 5. Nonarithmetic Operations; 6. Subroutines; 7. Input-Output; 8. Assembly of Complete Programs; 9. FORTRAN: Mixed Language Programs.

I agree with the authors that this book properly belongs at the fourth year level.

My only substantial criticism of this book is that it was written with the CDC 1604 digital computer as a reference. In fact, on occasion, the book reads like an instruction manual for this computer. I recognize that many ideas on languages can be explained only in the context of a specific structured computer. Still, I think the book could have been improved if the authors had indicated which comments applied to all machines and which related only to the CDC 1604. A second factor which detracts from the book's utility, particularly for the college student, is the almost total lack of a bibliography. The only reference I could find is imbedded in the text on page 230.

To summarize, this is an excellent book on the languages of the CDC 1604. If the reader is interested in another computer, he can derive much information from this book if he notes where the concepts apply to his computer.

LEON LEVINE, Scientific Data Systems

Problems in Higher Algebra. By D. K. Faddeev and I. S. Sominskii. Translated by J. L. Brenner. Freeman, San Francisco, 1965. 498 pp. $3.95.

This book consists of a compilation of approximately 1000 problems from the more classical parts of linear algebra and the theory of equations. Among the topics treated are the following: complex numbers, roots of unity, computation of determinants (some requiring mathematical induction), linear equations, matrices (a few of the matrix problems arose at the 1964 Gatlinburg Conference), inequalities involving matrices and determinants, divisibility of polynomials, interpolation, location of zeros of polynomials, symmetric functions, resultants and discriminants, applications of the Jordan canonical form. Many of the problems are of routine drill variety, but a significant percentage require imagination and ingenuity and will test the better students. The problems usually are independent of each other. Hints for the tougher problems are given in a special section and answers to all problems are provided in another section. It is obvious that this book will be a valuable source of problems for any teacher or examiner. It could be a useful adjunct text in courses treating such topics.

D. J. LEWIS, University of Michigan

Computational Methods of Linear Algebra. By D. K. Faddeev and V. N. Faddeeva. Translated by R. C. Williams. Freeman, San Francisco, 1963. xi+620 pp. $11.50.

This book is devoted to an exposition of computational methods for solving such basic problems of linear algebra as the solution of a system of linear equations, the inversion of a matrix, the determination of eigenvalues and eigenvectors of a matrix. There are a large number of numerical methods for solving these problems and the authors have grouped them about certain general points of view.

There are two basic methods of approach. An exact method is one that gives a solution to the problem by use of a finite number of elementary arithmetic operations. The number of operations necessary for a solution depends only on the computational scheme chosen and the order of the matrix defining the problem. If the coefficients are rational and the computations are carried out exactly, the solutions are exact. Historically, the first exact method was Gauss' method for solving a system of linear equations by eliminating the unknowns. There are many different computational schemes for applying Gauss' method and some of these are described in Chapter 2. An iterative method provides a means of determining an approximate solution, the solution being the limit of successive approximations obtained by iterating the same uniform process. Consequently, the rate of convergence plays an essential role in this process. No iterative method is universally effective in giving rapid convergence for all matrices. Usually the rule for the computational scheme for an iterative method is simple and convenient for actual computation. Iterative methods are seldom used in computing inverses of matrices. Various schemes for iterative solutions of a system of linear equations are given in Chapter 3.

The usual exact method of determining eigenvalues is the determination of the characteristic polynomial (with a minimum of computation) and then determination of the zeros of that polynomial. On the other hand, the iterative process usually enables one to avoid computing the characteristic polynomial. Consequently, an iterative process is easier and is used frequently when determining a single eigenvalue. An exact method is usually used when determining all the eigenvalues. Exact methods for the complete determination of eigenvalues are discussed in Chapters 4 and 6. The methods in Chapter 6 are based on the idea of orthogonalization. Iterative methods applicable to determining special eigenvalues are discussed in Chapters 5 and 7. Applications of the iterative method (via high speed computers) to the complete eigenvalue problem are discussed in Chapter 8. In Chapter 9, universal algorithms are discussed. A universal algorithm is an iterative process which is carried out according to formulas whose coefficients are the same for large classes of matrices.

Although this book is not encyclopedic it does contain a very large sample of the more useful algebraic methods applicable to these problems. Probabilistic methods, such as the Monte-Carlo methods, are not discussed. The various computational schemes are described with care, are theoretically justified, and are

illustrated with numerical examples which show how to set up the tabulation. A reader with a good acquaintanceship with linear algebra should be able to follow the discussion and apply the exact methods. He may need a small amount of elementary functional analysis to handle the iterative methods. The necessary background in linear algebra and in functional analysis is given in the opening chapter of 118 pages. The book concludes with an extensive list of references (59 pages) covering the years 1870–1959.

D. J. LEWIS, University of Michigan

Some Properties of Polyhedra in Euclidean Space. By V. J. D. Baston. Pergamon Press, Oxford, 1965. xi+212 pp. $10.00.

While the title of this book is faithful, it scarcely conveys the severe limitation of the book's contents. The entire volume, which is essentially the thesis for Baston's Ph.D. degree granted by the University of London, is devoted to a single, but very interesting, problem in 3-space. The problem arose from Crum's Problem, which in turn had its origin in the extension to 3-space of the famous Four Color Problem of the plane and sphere. Crum's Problem asks, "What is the maximum number of nonoverlapping convex polyhedra such that each pair has a common boundary of positive area?" It has been shown that there is no upper limit to this number, which proves that 3-space has no finite chromatic number. Now Crum's Problem has itself been extended in two ways: (1) by considering n-dimensional polyhedra in an n-dimensional space, and (2) by limiting the convex polyhedra in Crum's Problem to polyhedra of some special type, such as to tetrahedra or to cuboids. Baston's Ph.D. thesis problem was to investigate the question, "What is the maximum number of nonoverlapping tetrahedra such that each pair has a common boundary of positive area?" In a succession of 11 chapters, covering over 200 pages, Baston manages to show that the desired number is either 8 or 9. He actually constructs a set of 8 tetrahedra having the required property, but is unable to determine whether or not the number can be raised to 9.

The work, which is written on the research level, is a *tour de force* employing a matrix representation of the geometric problem. In the course of the work the reader is introduced to a vast array of new terms, such as *n-con*, *uno*, *unun*, *nodo*, *doun*, *notre*, *treun*, *dodo*, *dotre*, *tetre*, *naik*, *dhoat*, *thaik*, *daik*, *dopt*, *nopt*, *sipt*, *surplust*, *tredho*, *treno*, *tript*, and *chardho*. The analogue of the problem in 2-space is so very simple that one is surprised at the seeming complexity of the 3-dimensional case. Simpler methods of attacking the problem may be discovered in the future, but at present there is a lack of any developed theory to cope with the type of problem involved. Also, there is present that often observed discrepant gap between a 2-dimensional problem and its extension to 3-space. Thus, while the number sought in Crum's Problem in 3-space is nonfinite, the number for the 2-space analogue is easily shown to be the small finite number 4.

HOWARD EVES, University of Maine

Systems and Simulation. By D. M. Chorafas. Volume 14 of Mathematics in Science & Engineering Series. Academic Press, New York, 1965. 500 pp. $14.50.

This book is an introduction to the computer simulation approach to solving problems in engineering and operations research. The philosophy of properly formulating and evaluating management problems is discussed at length; here, of course, mathematics is a minor aspect. More than thirty different applications of simulation are then described, ranging from hydrology to model changes in the automobile industry. The discussion on each example tends to be rather general, and it might have been useful to show in more detail the difficulties that can arise. There is no bibliography.

J. J. FLORENTIN, Brown University

Error Propagation for Difference Methods. By Peter Henrici. Wiley, New York, 1963. vii+73 pp. $4.95.

This book is a sequel and companion to the author's earlier book [Discrete Variable Methods in Ordinary Differential Equations, Wiley, New York, 1962; this MONTHLY, 69 (1962) 935–936]. It extends the theory of linear multistep methods to systems of ordinary differential equations. However, the step from one dimension treated in the earlier book to several dimensions treated in the present book is not a mere exercise in easy generalization. Instead the treatment is essentially independent of the one-dimensional case and yields some new and unexpected results.

After a brief introduction in Chapter 1, the basic definitions and concepts are given in Chapter 2. The main result of Chapter 3 is the proof that stability and consistency are together necessary and sufficient for convergence. Chapter 4 presents a detailed study of the asymptotic behavior of the discretization error for small step sizes. In Chapter 5 round-off is considered and a statistical theory of the round-off error is given.

An appendix written with Riley and Stafford applies the results of Chapters 4 and 5 to the integration of two simple model problems each involving the integration of four simultaneous differential equations. Although the two solutions are the same, one problem is linear, while the other is not, and this leads to differences in error propagation between the two problems, and, in the nonlinear problem between different methods of integration. Experimental calculations which verify these results are summarized at the end of the appendix.

This little book is concise and carefully written. It should be of interest to all who are faced with the problem of integrating systems of ordinary differential equations. The reader who is willing to expend the effort necessary to read the book will be amply rewarded.

J. G. HERRIOT, Stanford University

Géométrie et relativité. By J.-M. Souriau. Hermann, Paris, 1964. 512 pp. 54 F.

The first part contains the fundamentals of geometry pertinent to the theory of relativity. With the aid of modern geometric methods the reader is introduced directly into the middle of differential geometry (variance theory, Riemannian geometry), matrix and tensor calculus, quaternions, Euclidean geometry, etc. In the second part there is a discussion of the fundamentals of the theory of relativity. Starting from a single variational principle the author derives the gravitational equation of Einstein and passes from general relativity to special, including derivations of the traditional concepts of space-time, and discussions of Newtonian mechanics and experimental verifications. The last part contains an exposition of the theory of relativity of five dimensions, elaborated in the past by Kaluza, Klein, Jordan, Pauli, and others, involving a unification of electricity with energy and impulse, light with gravitation, etc. A classical relativist will enjoy reading a book full of sophisticated geometric concepts and notions, regardless of how ideal the theory of relativity may be. But some remarks are in order. Fock demonstrated that the Einstein theory of gravitation is actually a development of the theory of Galilean space, not in the sense of a widening or generalization of the notion of relativity, but, on the contrary, in the sense of a limitation of this notion. This is ignored by the author. The present day controversy centering around the non-constancy of the speed of light, different times (biological, etc.) subject or not to the relativistic concept, influence of a neutrino flux upon the speed of light, time-dilatation dilemma, etc., is not mentioned.

M. Z. v. KRZYWOBLOCKI, Michigan State University

Principles of Random Walk. By Frank Spitzer. Van Nostrand, Princeton, 1964. 406 pp. $12.50.

This book is devoted almost exclusively to random walks on finite (mostly 1 or 2) dimensional lattices with spatially homogeneous transition probabilities. The most surprising thing about the book is that despite its length and high degree of specialization, it is written in such a style that it will hold the reader's interest even if this is not his own field. The main emphasis is on the interconnections among probability theory, harmonic analysis, and potential theory. By limiting the scope to walks on a lattice, the author can illustrate this very elegantly without the use of difficult topological concepts. He is obviously trying to teach and not to overpower the reader with mathematical magic. The book contains a thorough discussion of the first passage times of various points or sets from a fixed origin, and the mean number of occurrences of various events, for recurrent or transient walks. Most of the results are of recent origin but originally written in a form for experts. There are numerous examples and exercises.

G. F. NEWELL, Brown University

The Theory of Groups: An Introduction. By Joseph J. Rotman. Allyn and Bacon, Boston, 1965. xiii+305 pp. $9.50.

This is an excellent and also a very useful book. It is truly an introduction to group theory, avoiding, on the whole, very lengthy proofs and the buildup of elaborate machinery. But it is far from being elementary or primitive. Where the choice of topics was not a matter of necessity, the author has selected impressive results, frequently very recent ones, which may be considered as basic and characteristic for the whole complex of theorems to which they belong. The brief introduction to homological algebra and the presentation of the Novikov-Boone theorem on the unsolvability of the word problem will probably be particularly welcome to many readers.

Sections are preceded by lucid summaries. Motivation and clarity are stressed, on occasion even at the expense of brevity.

There are numerous exercises and examples. The prerequisites are kept at a minimum; aids from other parts of algebra are briefly developed or at least summarized.

List of chapters: Groups and homomorphisms. The isomorphism theorems. Permutation groups. Finite direct products. The Sylow theorems. Normal and subnormal series. Infinite abelian groups. Homological algebra. Free groups and free products. The word problem. Appendices.

W. MAGNUS, New York University

Counterexamples in Analysis. By B. R. Gelbaum and J. M. H. Olmsted. Holden-Day, San Francisco, 1964, 220 pp. $7.95.

This book will cause a stir in the pedagogical community as professors choose between encouraging students to buy this excellent and enjoyable book and banning it from library and bookstore. Now everyone has easy access to the classical and unusual counterexample which we can use so effectively in the classroom to bring home a point. It is nonetheless worth any inconvenience to have this book available. Since only one counterexample per assertion is presented, someone's favorite will be omitted; but at least there is one counterexample. The authors have performed a valuable service in compiling a fairly complete list with explanations and, where appropriate, references in the literature. The following list of chapter headings will indicate the coverage: 1. The real number system; 2. Functions and limits; 3. Differentiation; 4. Riemann Integration; 5. Sequences; 6. Infinite series; 7. Uniform convergence; 8. Sets and measure on the real axis; 9. Functions of two variables; 10. Plane sets; 11. Area; 12. Metric and topological spaces; 13. Function spaces.

The book seems to be unusually free from misprints, although the list of special symbols at the end omits reference to $A\backslash B$—the difference between sets A and B. The table of contents contains a full statement of each example, and seemed to be more useful than the index in locating particular examples. Such a book needs a kind of instant index for classroom emergencies. The theoretical

material is assumed known, but motivation is frequently supplied by careful and complete definitions of the terms, as well as by elucidating and elaborating remarks that go beyond the formal presentation of the example.

This book must have been fun to write because it was fun to read. It amused this reviewer to open the book at some random page and to challenge his honors seminar to come up with the required counterexample. The publishers also deserve kudos for publishing a book that is obviously not a textbook.

DAVID ROSEN, Swarthmore College

Absolute Stability of Regulator Systems. By M. A. Aizerman and F. R. Gantmacher. English translation by E. Polak. Holden-Day, San Francisco, 1964. 182 pp. $8.95.

The problem of the "absolute stability" was first formulated by Soviet engineers Lur'e and Postnikov in 1944. The regulator without control is linear and the feed-back control contains a single nonlinearity whose characteristic lies in the first and third quadrant. Absolute stability means asymptotic stability in the large of the origin (zero error in control) for all such nonlinear characteristics. The problem has been studied extensively by mathematicians and engineers, particularly in the Soviet Union and the United States. This book presents the main results up to 1963 and includes those obtained by the Liapunov method and by the frequency method. More recent work on the relation for this problem between the two methods is presented in an appendix.

J. P. LASALLE, Brown University

BRIEF MENTION

Theorie und Anwendung der unendlichen Reihen, Fünfte berichtigte Auflage. By Konrad Knopp. Springer-Verlag, Berlin, 1964. xii+582 pp. DM. 48,00.

Analysis and Synthesis of Linear Time-Variable Systems. By Allen A. Stubberud. U. of California Press, Berkeley, 1965. 108 pp. $4.75.

Based on a Ph.D. dissertation in engineering. Emphasis is on synthesis, especially of feedback systems. A knowledge of linear ordinary differential equations is assumed.

Mathematical Theory of Reliability. By Richard E. Barlow and Frank Proschan. Wiley, New York, 1965. 252 pp. $11.00.

A research monograph which presents a survey of probabilistic models useful in solving reliability problems.

The Philosophy of Science. By Peter Caws. Van Nostrand, Princeton, N. J., 1965, 352 pp. $6.75.

Introduction to Basic FORTRAN Programming and Numerical Methods. By William Prager. Blaisdell, Waltham, Mass., 1965. 204 pp. $6.50.

From lecture notes for a one-semester course for students of applied mathematics at Brown University.

Mathematics and Science, Last Essays. By Henri Poincaré. Translated by John W. Bolduc. Dover, New York, 1965. 121 pp. $1.25.

Darstellende Geometrie, Zweite verbesserte Auflage. By Fritz Rehbock. Springer-Verlag, Berlin, 1964. (First edition in 1956) 235 pp. DM. 29,00.

Dynamics-Particles, Rigid Bodies and Systems, Vol. I. By R. L. Halfman. Addison-Wesley, Reading, Mass., 1962. 379 pp. $8.50.

Dynamics-Systems, Variation of Methods, and Relativity, Vol. II. By R. L. Halfman. Addison-Wesley, Reading, Mass. 1962. 230 pp. $6.75.

Developments in Theoretical and Applied Mechanics, Vol. 2. Edited by W. A. Shaw. Pergamon, New York, 1965. 651 pp. $22.50.

Proceedings of the Second Southeastern Conference on Theoretical and Applied Mathematics, Georgia Institute of Technology, March, 1964.

Integral Equations, 2nd rev. ed. By S. G. Mikhlin. Pergamon, New York, 1964. 341 pp. $12.50.

Three Dimensional Dynamics, 2nd ed. By C. E. Easthope. Butterworth, Washington, D. C., 1964. 424 pp. $14.50.

The Groups of Order $2^n (n < or = 6)$. By Marshall Hall Jr. and J. K. Senior. Macmillan, New York, 1964. 225 pp. $15.00.

These tables give for the first time the complete list of the 267 groups of order 64.

An Introduction to the Mathematics of Servomechanisms. By J. L. Douce. Van Nostrand, Princeton, N. J., 1964. 240 pp. $4.50.

This is a textbook for students of control engineering. It assumes only an elementary knowledge of calculus and emphasizes a practical approach rather than mathematical rigour.

Computing and Data Processing Society of Canada, Proceedings of the Fourth Conference. Ottawa, May 1964. U. of Toronto Press, Toronto, 1964. 65 pp. $5.00.

The Basic Laws of Arithmetic. By Gottlob Frege. Translated and edited by M. Furth. U. of California Press, Berkeley, 1964. 142+42 pp. (editor's introduction) $5.00.

The editor introduces his translation of Frege's classic with an extensive discussion of its central ideas and their relevance to later developments.

Invariant Imbedding and Time-Dependent Transport Processes: Modern Analytic and Computational Methods in Science and Mathematics. Vol. II. By Bellman, Kagiwada, Kalaba and Prestrud. Elsevier, New York, 1964. 256 pp. $8.00.

A Rand Corporation Research Study.

The Concept of Matter in Descartes and Leibniz. By R. Catesby Taliaferro. U. of Notre Dame Press, Indiana, 1964. 33 pp. $2.00.

Die mathematischen Hilfsmittel des Physikers. 7th ed. By Erwin Madelung. Springer-Verlag, Berlin, 1964. 536 pp. DM 49, 70.

Analysis in Function Space. Edited by Martin and Segal. MIT Press, Cambridge, Mass., 1964. 218 pp. $6.00.

Probability and Information Theory with Applications to Radar. 2nd ed. by P. M. Woodward. Pergamon, New York, 1965, 136 pp. $5.00.

Discretus Calculus: A Variable-Metric Approach to Physical Theory. By Herbert S. Ingham. Philosophical Library, New York, 1964. 197 pp. $6.00.

Elementary Differential Equations, 6th ed. By Lyman M. Kells. McGraw-Hill, New York, 1965. 430 pp. $7.95.
Extensively revised.

Proceedings of the Conference on Complex Analysis, Minneapolis, December, 1964. A. Aeppli, E. Calabi and H. Röhrl, editors, Springer-Verlag, New York, 1965. 308 pp. $9.50.

Random Negative Exponential Deviates, No. 27, Tracts for Computers. By V. D. Barnett, edited by M. S. Bartlett. Cambridge U. Press, New York, 1965. 90 pp. $2.25.

Elements of Finite Probability. By J. L. Hodges and E. L. Lehmann. Holden Day, San Francisco, 1965. 230 pp. $5.25.
Part I of the author's book *Basic Concepts of Probability and Statistics.*

Cybernetics: or Control and Communication in the Animal and the Machine. 2nd ed. By Norbert Wiener. MIT Press, Cambridge, Mass., 1964. 212 pp. $1.95.

Coding Theorems of Information Theory, 2nd ed. By J. Wolfowitz. Springer-Verlag, New York, 1964. 156 pp. $6.75.

Methods of Applied Mathematics, 2nd ed. By Francis B. Hildebrand. Prentice-Hall, Englewood Cliffs, N. J., 1965. 360 pp. $10.50.

An Introduction to the Foundations and Fundamental Concepts of Mathematics, rev. ed. By Howard Eves and Carroll V. Newsom. Holt, Rinehart and Winston, New York, 1965. 414 pp. $7.75.

Statistical Theory and Methodology in Science and Engineering, 2nd ed. By K. A. Brownlee. Wiley, New York, 1965. 590 pp. $18.50.

Regulare Figuren. By L. Fejes Tóth. Akademiai Kiado, Budapest, 1965. 316 pp. $8.50.
A review of the English translation will appear in the MONTHLY.

The Treasury of Mathematics. Henrietta O. Midonick, editor. Philosophical Library, New York, 1965. 820 pp. $15.00.

Principles of Logic. By Josiah Royce. Wisdom Library, Philosophical Library, New York, 1961. 77 pp. $2.75.
A representative sampling of Royce's logical essays.

Nomography. By Edward Otto. Pergamon, New York, 1963. 313 pp. $10.00.
This book is a manual for teaching nomography. It stresses the importance of geometrical transformations, especially projective transformations of the plane.

Les Anneaux Normes Commutatifs. By I. M. Gelfand, D. A. Raikov, and G. E. Šilov. Gauthier-Villars, Paris, 1964. 258 pp. 52 F.
The English translation of this work will be reviewed.

Principles of Coding, Filtering, and Information Theory. By Leonard S. Schwartz. Spartan Books, Baltimore, 1963. xiii+255 pp. $8.50.

Pedal Lines, Orthopoles, Kantor Lines, Kantor Points, Deltoids and Poristic Polygons. (Study by the Method of Complex Coordinates). By Kesiraju Satyanarayana. Privately printed, Rajahmundry, Andhra Pradesh, India, 1964. 148+xii pp. R. 3.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor E. M. Beesley, University of Nevada, represented the Association at the inauguration of Newton E. Miller as First Chancellor of the Reno Campus of the University of Nevada on February 11, 1966.

Professor J. N. Eastham, Queensborough Community College, represented the Association at the installation of Dr. Allyn P. Robinson as Dean of Adelphi Suffolk College on March 6, 1966.

Professor Mariano Garcia, University of Puerto Rico, represented the Association at the inauguration of Dr. Raymond B. Hoxeng as President of Inter-American University of Puerto Rico on March 6, 1966.

Mr. M. N. Chase, formerly Director of the Combat Operations Research Group of Technical Operations Research, Arlington, Virginia, has been appointed Vice President of the Washington Area Activities.

Dr. A. S. Householder, Oak Ridge National Laboratory, was awarded the degree Doktor der Naturwissenschaften Ehrenhalber by the Technische Hochschule, Munich, Germany, on December 2, 1965.

Associate Professor J. N. Javaher, Stanislaus State College, has been promoted to Professor.

Mr. J. C. Lanz, Harrisburg Area Community College, has been promoted to Professor.

Mr. John A. Lushbough, Colorado State University, has been appointed to the staff of the Oak Ridge National Laboratory as a Programmer.

Assistant Professor L. V. Quintas, St. John's University, has been awarded an NSF Science Faculty Fellowship for a twelve month period of study in mathematics at the City University of New York.

Associate Professor D. E. Varberg, Hamline University, has been promoted to professor.

Professor Jesse Douglas, City University of New York, died on October 7, 1965. He was a member of the Association for 22 years.

Professor Emeritus E. B. Stouffer, University of Kansas, died on November 24, 1965. He was a charter member of the Association.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

"FINDING EMPLOYMENT IN THE MATHEMATICAL SCIENCES" BROCHURE AVAILABLE

The Mathematical Sciences Employment Register, which is sponsored by the American Mathematical Society, the Mathematical Association of America, and the Society for Industrial and Applied Mathematics, has recently published a booklet entitled "Finding Employment in the Mathematical Sciences." This booklet gives information to the

young mathematician, who is just entering the professional field, information on how to find employment best suited to his abilities and training. The role of the mathematician in teaching, academic and industrial research, computing, and government is discussed. The booklet, also, lists the sources of information available to the young mathematician who is seeking a position. There is no charge for this booklet, and it may be obtained by writing to the Employment Register, P.O. Box 6248, Providence, Rhode Island 02904.

MATHEMATICAL SCIENCES EMPLOYMENT REGISTER

The Mathematical Sciences Employment Register, established by the American Mathematical Society, the Mathematical Association of America, and the Society for Industrial and Applied Mathematics, will be maintained at the Summer Meeting at Rutgers—The State University, New Brunswick, New Jersey on August 30–31 and September 1, 1966. The Register will be open from 9:00 a.m. to 5:00 p.m. on each of these three days in Frelingheusen Hall, Level A. It is important that applicants and employers register at the Employment Register Desk promptly upon arrival at the meeting to facilitate the arrangement of appointments.

There is no charge for registration, either to job applicants or to employers, except when the late registration fee for employers is applicable. Provision will be made for anonymity of applicants upon request and upon payment of $5.00 to defray the cost involved in handling anonymous listings.

Job applicants and employers who wish to be listed should write to the Employment Register, P.O. Box 6248, Providence, Rhode Island 02904, for applicants forms or for position description forms. These forms must be completed and returned to Providence not later than July 15, 1966, in order to be included in the listings at the Summer Meeting in New Brunswick. Position description forms arriving after this closing date will be charged $5.00 for a late listing. Those forms which arrive too late to be included in the printed list are taken to the meeting where they may be seen by interested applicants. The printed lists will be mailed to subscribers on August 5, 1966 and will be available for distribution both during and after the meeting.

A subscription which includes all three issues (January, May, and August) of both the list of applicants and the list of positions is obtainable for $15 per year. The individual issues of both lists may be purchased in January, May, and August for $7.50. Copies of only the list of positions may be purchased for $3.00.

Persons wishing either a year's subscription or individual issues, should make checks payable to the American Mathematical Society and send them to the Mathematical Sciences Employment Register, P.O. Box 6248, Providence, Rhode Island 02904.

MATHEMATICS FILMS

Many of the Individual Lectures films produced by the MAA Committee on Educational Media are now available for purchase or rental. Experimental versions of some of these films have been a feature of national MAA meetings for the past two years. Readers wishing to order such films, or further information, are urged to write *not* to the MAA, but to: Modern Learning Aids, 1212 Avenue of the Americas, New York, N. Y. 10036.

MATHEMATICS ON TELEVISION

Would persons who are aware of mathematical broadcasts on television, open or closed circuit, please drop a note to Phillip S. Jones, Mathematics Department, University of Michigan, Ann Arbor, Mich. 48104, chairman of a subcommittee of the MAA Committee on Educational Media.

NEW MAA PUBLICATION: HARMONIC ANALYSIS

MAA is publishing a single volume edition of "Harmonic Analysis" by Lynn H. Loomis. This book of about 400 pages consists of notes by Ethan Bolker of lectures given by Professor Loomis at the 1965 MAA Cooperative Summer Seminar at Bowdoin College.

In the preface the author states: "The principal theme of these notes is the development of the heavy machinery required to manipulate the Fourier Transform. Secondary themes are the consistent applications of norms of increasing diversity and complexity and the gradual introduction of weak methods, starting with the pairings of the classical Banach spaces and culminating with the introduction of distributions. . . . The nominal aim of the lectures is to prove in a modern way a few theorems about linear partial differential equations."

Copies will be ready about June 1, and may be purchased for $3.00 from: MAA, SUNY at Buffalo, Buffalo, N. Y., 14214. Payment is requested with your order.

CALENDAR OF FUTURE MEETINGS

Forty-seventh Summer Meeting, Rutgers, The State University, New Brunswick, New Jersey, August 29–September 1, 1966.

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

Allegheny Mountain
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana-Mississippi, Jung Hotel, New Orleans, March 4–5, 1967.
Maryland-District of Columbia-Virginia
Metropolitan New York
Michigan
Minnesota
Missouri
Nebraska
New Jersey
Northeastern
Northern California, University of California, Davis, February 4, 1967.
Ohio
Oklahoma-Arkansas
Pacific Northwest, University of Victoria, Victoria, British Columbia, June 17, 1966.
Philadelphia, Villanova University, Villanova, November 1966.
Rocky Mountain
Southeastern, Florida Presbyterian College, St. Petersburg, Spring, 1967.
Southern California, San Diego State College, San Diego, March 18, 1967.
Southwestern
Texas
Upper New York State
Wisconsin

FUTURE MEETINGS OF OTHER ORGANIZATIONS

American Association for the Advancement of Science, Washington, D. C., December 26–31, 1966.

American Mathematical Society, Rutgers, The State University, New Brunswick, N. J., August 30–September 2, 1966.

American Society for Engineering Education, Washington State University, Pullman, June 20–24, 1966.

Association for Computing Machinery, Ambassador Hotel, Los Angeles, August 30–September 1, 1966.

Central Association of Science and Mathematics Teachers, Indianapolis, November 24–26, 1966.

Institute of Mathematical Statistics, Rutgers, The State University, New Brunswick, New Jersey, August 30–September 2, 1966.

National Council of Teachers of Mathematics, Miami Beach, Florida, June 29, 1966.

Operations Research Society of America, Jack Tar Hotel, Durham, N. C., October 17–19, 1966.

Society for Industrial and Applied Mathematics, State University of New York, Stony Brook, September 12–14, 1966.

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 — NUMBER 6

CONTENTS

JUNE-JULY 1966

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453. Manuscripts of papers should be submitted to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE TO AUTHORS

The MONTHLY welcomes papers presenting valid mathematics, of rather general interest, at a level intelligible to persons with two years of full-time graduate study. Some novelty of content, viewpoint, or arrangement is essential. Expository articles are particularly desired. State the context and the principal aim of the paper early. Address yourself quite explicitly to the reader described above, communicating your ideas to him clearly and attractively.

The title should be brief and meaningful. Since the title will be quoted and reproduced by laymen, it should contain no symbols unfamiliar to laymen.

Articles should be typewritten, double-spaced, on $8\frac{1}{2}\times 11''$ paper of very good quality. Submit the original (and a duplicate if convenient) keeping a complete copy for yourself. To avoid loss and delay notify us of any change of address.

The typescript should be prepared with extreme care. Misprints are highly obnoxious; so are dangling participles. Put name and address between title and text. Put references at the end with bracketed citations in the text. Avoid footnotes; instead, use clearly designated remarks in the text. Put acknowledgments at the very end, just before the bibliography. Be generous with spacing and displays. *Keep notation simple.* For a matrix the notation $[a_{jk}]$ is recommended, with $\det[a_{jk}]$ or $|a_{jk}|$ for the determinant. On doubtful questions regarding format or notation, observe practices in current issues of the MONTHLY, or consult the applicable sections of the "Author's Manual" of the American Mathematical Society.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y.
during the months of January, February, March, April, May, June–July,
August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

METRIZATION

F. BURTON JONES, University of California, Riverside

In the 1910's and 20's there was considerable interest in finding topological characterizations of metric spaces (see Chittenden's summary in [4]). Some of these were the result of unsuccessful attempts to obtain topological generalizations of metric spaces [**7**, **8**, **15**]. R. L. Moore was successful in obtaining such a generalization but in the process he also discovered a characterization but never published it. While this theorem has been referred to by various authors and even Moore pointed to its existence, I think that he probably felt it to be too simple a consequence of Chittenden's theorem about the equivalence of *écart* and *voisinage* to warrant publication. Nevertheless, some of the rather simple consequences of Moore's theorem do not appear to be so simple after all and are being rediscovered by Russian mathematicians. Hence, it seems to me to be worthwhile to discuss in some detail this little-known metrization theorem and some of its consequences.

1. Moore's Metrization Theorem. *Let S be a topological space without contiguous points (i.e., a T_1-space). Suppose that there exists a countable family $\mathfrak{F}$ of open covers of S such that if H is a closed subset of S and p is a point of S-H, there is a cover F in $\mathfrak{F}$ such that no element of F containing p intersects any element of F containing a point of H (i.e., $F^*(p)\cap F^*(H)=\varnothing$). Then S is metrizable* [p. 21 in **11**, **14**, **16**, **17**].

While this is not precisely the form in which Moore stated the theorem, I am stating it in this form to emphasize the kind of uniform regularity (or normality) contained in the hypothesis that the sets, $F^*(p)$ and $F^*(H)$, are disjoint open subsets of S containing p and H respectively. It is also easy to see why one would expect the theorem to be true. For suppose that S is metric and that F_n is the collection of all open subsets of S of diameter $1/n$ or less ($n=1, 2, 3, \cdots$). Clearly F_n covers S and if $1/n<d(p, H)/2$, then $F^*(p)\cap F^*(H)=\varnothing$ by the triangle inequality for the metric distance function d. Furthermore, one cannot strengthen the hypothesis (in the obvious way) by requiring for disjoint closed sets H and K that $F^*(H)\cap F^*(K)=\varnothing$ because, as the reader may verify, this is not true for some of the simplest metric spaces (e.g., the plane) no matter how $\mathfrak{F}$ is defined. Of course, if $d(H, K)$ were positive, the above argument *would* work and this would be the case if either H or K were compact. But for this situation a slightly stronger form of Moore's Theorem is possible.

2. The strong form of Moore's Metrization Theorem. *Let S be a regular topological space ($=$regular Hausdorff space). Suppose that there exists a countable family $\mathfrak{F}$ of open covers of S such that if H and K are disjoint closed subsets of S, one of which is compact, there is an element F of $\mathfrak{F}$ such that no element of F intersects both H and K (i.e., $F^*(H)\cap K=\varnothing=F^*(K)\cap H$). Then S is metrizable.*

This follows quickly (and indirectly) from Moore's Metrization Theorem as follows:

Let $F_1, F_2, F_3, \cdots$ be a simple well-ordering of $\mathfrak{F}$ and let $G_1, G_2, G_3, \cdots$ be a sequence of open covers of S such that for each natural number n, G_n is a common refinement of $F_1, F_2, \cdots, F_n$ (i.e., each element of G_n is a subset of some element of F_i for $1 \leqq i \leqq n$). Suppose that S is not metrizable; then the countable family $\{G_n\}$ must fail to satisfy the hypothesis of Section 1 when $\{G_n\}$ is substituted for $\mathfrak{F}$. Hence there exist a closed set H and a point p of S-H such that for each n, $G_n^*(p) \cap G_n^*(H) \neq \varnothing$.

For each n, let p_n denote a point in this intersection and let K denote the closure of the set of all these points. Since $\{p\}$ is closed and compact, the sequence $p_1, p_2, p_3, \cdots$ converges to p. Hence H and K-H are disjoint closed sets and K-H is compact; so there must exist a natural number n such that no element of F_n intersects both. Hence the same must be true of each of $G_n, G_{n+1}, G_{n+2}, \cdots$, which is a contradiction because for i large enough for p_i to belong to K-H, $G_i^*(H)$ obviously intersects both H and K-H. The theorem follows from this contradiction.

3. Moore Spaces. In the introduction mention was made of the search for topological generalizations of metric spaces. The generalization discovered by Moore [**13**] is as follows (compare this with Section 2):

Let S be a regular topological space (=regular Hausdorff space). Suppose that there exists a countable family $\mathfrak{F}$ of open covers of S such that if H is a closed subset of S and p is a point of S-H, then there is a cover F in $\mathfrak{F}$ such that no element of F contains p and intersects H. Such a space is called a Moore space and $\mathfrak{F}$ is called a development for the space. Quite clearly all metric spaces are Moore spaces, for as before, if F_n is the collection of all open subsets of a given metric space of diameter $1/n$ or less ($n=1, 2, 3, \cdots$), F_n covers the space and the family $\{F_n\}$ has all of the properties required of $\mathfrak{F}$. And it is easy to speculate that perhaps all Moore spaces are metric spaces.

It is well known that in a great many situations what is true of closed and finite point sets is true of closed and compact point sets and conversely. So let us restate Moore's strong metrization theorem of Section 2, substituting "finite" for "compact" as follows (again let S denote a regular Hausdorff space):

Suppose that there exists a countable family $\mathfrak{F}$ of open covers of S such that if H and K are disjoint closed subsets of S, one of which is finite, then there is an element F of $\mathfrak{F}$ such that no element of F intersects both H and K (i.e., $F^*(H) \cap K = \varnothing = F^*(K) \cap H$). Is every space of this sort metric (=metrizable)?

The reader can readily verify that every Moore space is such a space for it is no loss of generality to suppose that the development $\mathfrak{F}$ postulated to exist for a Moore space has a simple well-ordering, $F_1, F_2, F_3, \cdots$ and that F_n refines F_i for $i<n$. Hence, K being a Moore space, for each point p of K there exists a cover F in $\mathfrak{F}$ such that $F^*(p) \cap H = \varnothing$. When K is finite this set of covers is finite and which ever one has the largest subscript (in the well-ordering) has

the required property—namely, no element of it intersects both H and K. Thus, one may see that every Moore space is a space of this (the above) sort. But not every Moore space is metric; so not every space of this sort is metric despite the strength of the "compact = finite" principle.

4. Nonmetrizable Moore Spaces. The best-known example of a Moore space which is not metrizable is due to R. L. Moore. Let S denote the points of the plane on or above the x-axis and let regions (= basis elements) be of two types: (1) if a circle lies entirely above the x-axis its interior is a region and (2) if a circle is tangent to the x-axis (from above) then its interior plus the point of tangency is a region. Let F_n ($n=1, 2, 3, \cdots$) be the collection of all regions of diameter $1/n$ (here I am using the ordinary Euclidean distance). The family $\{F_n\}$ has the properties demanded of $\mathfrak{F}$ to make S a Moore space. Since no region contains two points of S belonging to the x-axis, the x-axis has become a discrete point set. If S were metrizable, S would necessarily have a countable basis (the part of S above the x-axis is still separable just as it was in the plane) and every *countable* basis would leave uncountably many points on the x-axis uncovered. Hence S is not metrizable.

Moore has in [**12**] another example of this sort (but harder to describe) which is connected (in fact, cyclicly connected), locally planar (basis elements are homeomorphic with open plane disks), globally planar (the Jordan curve theorem holds true), separable, but nonmetrizable. What is probably the most complicated and most nonmetric known example of a Moore space is due to M. E. Estill (Rudin) [**6**]. This Moore space has the property that every metric subspace is nowhere dense.

5. Metrization of Moore Spaces. Since a Moore space is not necessarily metrizable, the question arises as to just how a Moore space differs from a metric space. Many years ago I thought that the property of *normality* was the only difference (and this is indeed the case if some countable set is dense in the space [**10**]). But despite the efforts of many investigators this problem is still unresolved. At present the best answer to the question is *paracompactness*. That every paracompact Moore space is metric was first observed by Bing [**3**]. I give below a somewhat different argument.

Suppose that S is a paracompact Moore space and that $\{F_n\}$ is a development for S. Since S is regular let us assume that $\{F_n\}$ has the property that not only does F_{n+1} refine F_n but that the collection of closures of elements of F_{n+1} does also (i.e., if f belongs to F_{n+1} then $\bar{f}$ is a subset of at least one element of F_n). Since S is paracompact let us also assume that for each natural number n, F_n is locally finite (i.e., if p is a point, there is an open set U which contains p and intersects only finitely many elements of F_n). Let H and K denote disjoint closed subsets of S such that one of them, say K, is compact and for each n, let $F_n(K)$ denote the collection of all elements of F_n which intersect K. Since K is compact and F_n is locally finite, $F_n(K)$ is finite. Now if for each n, $F_n^*(K)\cap H \neq \varnothing$, there exists a sequence $f_1, f_2, f_3, \cdots$ such that $f_1 \supset \bar{f}_2 \supset f_2 \supset \bar{f}_3 \supset f_3 \supset \cdots$

and for each n, f_n is an element of $F_n(K)$ which intersects both H and K. Let p be a point of $\bigcap(K\cdot\bar{f}_n)$. Then p is a point of K which for each n belongs to f_n. So for each n, $F_n^*(p)\cap H\neq\varnothing$ contrary to the required behavior of a development for the Moore space S. Consequently $\{F_n\}$ has the properties of $\mathfrak{F}$ required in Section 2 to make S metrizable.

6. Spaces with a uniform base. Alexandroff recently introduced (in [1]) the notion of a space having a uniform base (not to be confused with a space having a uniform structure). A base B is uniform if and only if each infinite subcollection of $B(p)$ is a base at p for each point p in the space. Every metric space has a uniform base because paracompactness guarantees that for each natural number n some locally finite subcollection of all open sets of diameter less than $1/n$ covers the space. On the other hand, not every space having a uniform base (even when the space is also regular) is metrizable. The nonmetrizable Moore space of Section 4 does not have a uniform base despite the fact that the complement of the x-axis has the topology of the plane. However, something of a modification of it [due to R. W. Heath] does have a uniform base without becoming metric.

Let S denote the points of the plane on or above the x-axis and let regions (=basis elements) be of two types: (1) if p is a point above the x-axis then $\{p\}$ is a region (i.e., the portion above the x-axis is given the discrete topology) and (2) if a circle of radius one is tangent to the x-axis at p (from above) then an open interval of it (with p as the mid-point) is a region. Let F_n ($n=1, 2, 3, \cdots$) be the collection of all regions of diameter $1/n$ or zero (again I am using the ordinary Euclidean distance). The family $\{F_n\}$ is a development and S is a Moore space. Clearly $\bigcup F_n$ is a uniform base for S. But S is not metrizable since S is not normal. (The reader can quickly check this by observing that since the set H of all irrational numbers on the x-axis and the set K of all rational numbers are closed and disjoint, normality would require a cover Q of H by basis elements such that Q^* has no point or limit point in K. This means that the real function f which assigns zero to each rational number and the diameter of the element of Q containing it to each irrational number is actually continuous at each rational number and discontinuous at each irrational number. But no such function f exists.)

These spaces (those with uniform bases) have a very simple relation to Moore spaces. Specifically, in order that a regular Hausdorff space S have a uniform base it is necessary and sufficient that S be a point-wise paracompact Moore space. (This was called to my attention by Heath.) For suppose that $\mathfrak{F}$ is a development for S. As before, one loses no generality in assuming that $\mathfrak{F}=\{F_n\}$ where for each n, F_{n+1} refines F_n. Now if S is assumed to be point-wise paracompact each of these covers F_n has a point-finite refinement F_n' which covers S (i.e., no point belongs to infinitely many elements of F_n'). Hence $B=\bigcup F_n'$ is a uniform base.

Conversely, suppose that the regular Hausdorff space has a uniform base B. Let $B_1, B_2, \cdots, B_z, \cdots$ be a well-ordering of B. The cover F_1 is defined by

induction as follows: let the first element of F_1 be B_1 and each subsequent element of F_1 is the first element of $B_1, B_2, \cdots$ to contain a point not in any preceding element of F_1. This process produces a point-finite cover of S. In fact, let p be a point of S and let f be the first element of F_1 containing p. Since each element of F_1 which follows f in F_1 contains some point of S-f, the collection of all such elements which contain p does not form a base at p and (hence) is finite.

Now F_2 is defined as follows. Let B^1 be the subcollection of B remaining after removing from B all of the nondegenerate terms of F_1. Let the first term of F_2 be the first term of B^1 (in the well-ordering of B) and each subsequent element of F_2 is the first element of B^1 (in the well-ordering of B) to contain a point not in any preceding element of F_2. This process may be continued in exactly this same fashion (letting $B^0 = B$ and being certain not to delete the degenerate elements of F_i in B^{i-1} in producing B^i) to define $\{F_n\}$ which is a development $\mathfrak{F}$ for S.

To see that S is point-wise paracompact let G be an arbitrary open cover of S, let B' be a subcollection of B which refines G and covers S and let F' be selected from B' in exactly the same way that F_1 was selected from B. Then F' is point-finite.

So every regular Hausdorff space which has a uniform base is a point-wise paracompact Moore space.

Here again we may ask how such a space differs from a metric space. Alexandroff gives three answers (e.g., paracompactness) but these answers add nothing new to our knowledge because of the above connection with Moore spaces. And again we may ask:

Is every normal Hausdorff space which has a uniform base metrizable?

References

1. P. S. Aleksandrov, Some results in the theory of topological spaces, obtained within the last twenty-five years, Russian Math. Surveys, 15, (No. 2) (1960) 23–83.

2. A. Arhangelskii, Bicompact sets and the topology of spaces, Dokl. Akad. Nauk SSSR, 150 (1963) 9–12.

3. R. H. Bing, Metrization of topological spaces, Canadian J. Math., 3 (1951) 175–186.

4. E. W. Chittenden, On the metrization problem and related problems in the theory of abstract sets, Bull. Amer. Math. Soc., 33 (1927) 13–34.

5. ———, On the equivalence of écart and voisinage, Trans. Amer. Math. Soc., 18 (1917) 161–166.

6. Mary Ellen Estill, Concerning abstract spaces, Duke Math. J., 17 (1950) 317–327.

7. M. Fréchet, Sur les classes V normales, Trans. Amer. Math. Soc., 14 (1913) 320–324.

8. E. R. Hedrick, On properties of a domain for which any derived set is closed, Trans. Amer. Math. Soc., 12 (1911) 285–294.

9. F. B. Jones, R. L. Moore's Axiom 1′ and Metrization, Proc. Amer. Math. Soc., 9 (1958) 487.

10. ———, Concerning normal and completely normal spaces, Bull. Amer. Math. Soc., 43 (1937) 671–677.

11. R. L. Moore, A set of axioms for plane analysis situs, Fundamenta Mathematicae, 25 (1935) 13–28.

12. ———, Concerning separability, Proc. Nat. Acad. Sci. U.S.A., 28 (1942) 56–58.

13. ———, Foundations of Point Set Theory, Amer. Math. Soc. Colloquium Publ., 13 (1932) Rev. ed. 1962.

14. V. W. Niemytski, On the "third axiom of metric space," Trans. Amer. Math. Soc., 29 (1927) 507–513.

15. A. D. Pitcher and E. W. Chittenden, On the foundations of the calcul fonctionnel of Fréchet, Trans. Amer. Math. Soc., 19 (1918) 66–78.

16. C. W. Vickery, Axioms for Moore spaces and metric spaces, Bull. Amer. Math. Soc., 45 (1940) 560–564.

17. W. A. Wilson, On semi-metric spaces, Amer. J. Math., 53 (1931) 361–373.

18. R. W. Heath, Screenability, pointwise paracompactness and metrization of Moore spaces, Canad. J. Math., 16(1964) 763–770.

"DECIMAL" EXPANSIONS TO NONINTEGRAL BASES

L. C. EGGAN, Pacific Lutheran University AND
C. L. VANDEN EYNDEN, Miami University

1. Introduction. It is easy to see that for $\alpha>1$ we can associate with each number $\xi\geqq 0$ a "quasi-decimal" (cf. [1]) expansion to the base α of the form

$$\xi = \sum_{n=-\infty}^{\infty} a_n\alpha^{-n}, \tag{1}$$

where the a_n are integers such that $a_{-n}=0$ for n sufficiently large, and

$$0 \leqq a_n < \alpha. \tag{2}$$

For example, if $\xi<1$, the sequence of a_n defined recursively by

$$\begin{aligned} \xi_0 &= \xi, \qquad a_n = [\alpha\xi_{n-1}], \\ \xi_n &= \alpha\xi_{n-1} - a_n, \end{aligned} \tag{3}$$

satisfies (2), and

$$\xi = \sum_{n=1}^{\infty} a_n\alpha^{-n}.$$

For simplicity, we shall write the last equation as $\xi=.a_1a_2\cdots$, the base α being understood. If $\alpha=5/2$ and $\xi=6/7$ the algorithm defined by (3) gives us

$$6/7 = .200201\cdots. \tag{4}$$

By summing the geometric series, however, we note that also $6/7=.121212\cdots$. Thus (2) is not a sufficient condition for the expansion (1) to be unique.

If the sequence of a_n is repeating except for some initial set of terms we shall call the expansion *periodic*. The periodic expansion $.a_1\cdots a_nb_1\cdots b_mb_1\cdots b_m\cdots$ will be abbreviated $.a_1\cdots a_n(b_1\cdots b_m)$; thus $6/7=.(12)$ to base 5/2.

If α is an integer the theory of the above expansions is well known; *in what follows we shall assume, therefore, that α is greater than* 1 *and not an integer.*

In Section 2 we prove some simple preliminary results. In Section 3 we show that if α is rational, any number has at most one periodic expansion to base α. Thus the algorithmic expansion given in (4) above is not periodic, even though α and ξ are both rational. In Section 4 we obtain some results on the number of expansions that a number can have and about the special role of the algorithmic expansion. Finally, in Section 5 we obtain necessary and sufficient conditions for a number to have a periodic expansion to a rational base and an algorithm for finding this expansion.

2. Preliminaries. Without loss of generality we will restrict our attention to expansions in which $a_n=0$ for $n\leqq 0$. Recall that $1<\alpha\neq[\alpha]$ by assumption. Set $b=[\alpha]$.

LEMMA 1. *For any α, $.1<.0(b)$.*

Proof. $.0(b)=b\alpha^{-2}/(1-\alpha^{-1})>(\alpha-1)/(\alpha^2-\alpha)=\alpha^{-1}=.1$.

COROLLARY. *For any α, there exist intervals in which each number has more than one expansion.*

For example, each number in the interval $[.1, .0(b)]$ has at least two representations, one starting with 0 and one starting with 1.

LEMMA 2. *If a number has two different expansions to base α, then they differ first by exactly* 1.

Proof. If $\alpha<2$, then $b=1$ and the lemma is trivial. Suppose that $\alpha>2$. If two expansions could differ first by more than 1 we would clearly have $.2\leqq .0(b)$. This says $2\alpha^{-1}\leqq b\alpha^{-2}/(1-\alpha^{-1})$, or $2(1-\alpha^{-1})\leqq b\alpha^{-1}$. But then $1=2(1-2^{-1})<2(1-\alpha^{-1})\leqq b\alpha^{-1}=[\alpha]\alpha^{-1}<1$, a contradiction.

3. The number of periodic expansions. In this section we consider the number of periodic expansions a number can have to a given base. If $\alpha=(1+\sqrt{5})/2$, then $.1=.011=.01011=\cdots$, so that α^{-1} has infinitely many periodic (and, in fact, finite) expansions. For rational α the situation is different.

THEOREM 1. *If α is rational, then any number has at most one periodic expansion to base α.*

Proof. Let $\alpha=P/Q$, where $P>Q>1$ and $(P, Q)=1$. Suppose ξ has two periodic expansions. By subtracting them we obtain a periodic representation of zero,

$$0=\sum_{j=1}^{m} c_j\alpha^{-j}+\alpha^{-m}\sum_{j=1}^{k} d_j\alpha^{-j}(1+\alpha^{-k}+\alpha^{-2k}+\cdots), \tag{5}$$

where $|c_j|$, $|d_j|\leqq b$ for all j, some c or d is not zero, and m is minimal. If k is

zero, we have

$$0 = \sum_{j=1}^{m} c_j Q^j P^{m-j},$$

from which we see that P divides c_m. But $|c_m| \leqq b = [P/Q] < P/2$, and so $c_m = 0$, contradicting the minimality of m. Similarly $m \neq 0$.

Multiplying (5) by $P^m(P^k - Q^k)$ yields

$$0 = (P^k - Q^k) \sum_{j=1}^{m} c_j Q^j P^{m-j} + Q^m \sum_{j=1}^{k} d_j Q^j P^{k-j},$$

from which we see that $0 \equiv -c_m Q^{k+m} + Q^{m+k} d_k \pmod P$, or $c_m \equiv d_k \pmod P$. However, $|d_k - c_m| < P$ so that $c_m = d_k$, again contradicting the minimality of m. This proves the theorem.

Theorem 1 shows that the algorithmic expansion in (4) is not periodic, even though α and ξ are both rational. In fact, we have the following result.

COROLLARY. *For any rational α there exists a real number ξ such that ξ has a periodic expansion, but the expansion given by the algorithm is not periodic.*

Proof. For any rational α, $\xi = .0(b)$ is such a number. For $\alpha\xi = b/(\alpha - 1) > 1$, and so the algorithmic expansion starts with 1. By the theorem, this expansion cannot be periodic.

Note that once we have one number with this property, it is easy to construct infinitely many.

4. Results for general α. In this section we characterize those bases which admit numbers with unique expansions. We also investigate the algorithmic expansion.

THEOREM 2. (a) *If $\alpha > (1+\sqrt{5})/2$, then there exist infinitely many numbers with a unique expansion to base α.* (b) *If $\alpha \leqq (1+\sqrt{5})/2$, every positive number has infinitely many expansions to base α.*

Proof. (a) Suppose $\alpha > (1+\sqrt{5})/2$.

CASE I. $\alpha < 2$. Here $b = 1$. We will show that $\xi = .(01)$ has a unique expansion. Suppose $\xi = .a_1a_2 \cdots$ is a different expansion. The inequality $\xi < .1$ shows $a_1 \neq 1$; suppose $a_1 = a_2 = 0$. Then $\xi = .00a_3 \cdots \leqq .00(1)$. Subtracting $.00(01)$ yields $.01 \leqq .0(001)$, or $\xi \geqq .1$.

CASE II. $\alpha > 2$. Here $b \geqq 2$. Let $\xi = .(1)$ and suppose it has another expansion $.a_1a_2 \cdots$. We may as well assume $a_1 \neq 1$. If $a_1 = 0$, then $.0(b-1) > .1$, while if $a_1 = 2$, then $.0(1) > .1$. But

$$.0(1) \leqq .0(b-1) = \alpha^{-1}(b-1)/(\alpha - 1) < \alpha^{-1} = .1.$$

In both cases it is easily seen that $\alpha^{-n}\xi$ also has a unique expansion for any integer $n \geqq 1$.

(b) Suppose $\alpha \leqq (1+\sqrt{5})/2$. We note that $b=1$ and define ξ^* to be $.(1)-\xi$. Since interchanging the ones and zeros in any expansion of ξ gives an expansion of ξ^*; we see that ξ and ξ^* have the same number of expansions.

CASE I. Any expansion of ξ is eventually constant.

Suppose $\xi = .a_1 \cdots a_t 1$. Then, by Lemma 1, $\xi = .a_1 \cdots a_t 0 b_1 b_2 \cdots$, where the b's are not all zero. Thus ξ has an expansion where the constant part starts later. If an expansion of ξ is eventually 1, consideration of ξ^* gives the same result. Thus ξ has expansions where the constant part starts arbitrarily far out, so there must be infinitely many of them.

CASE II. ξ has an expansion containing 100 or 011 infinitely often.

Suppose there are infinitely many 100's. Since $.011 \geqq .100$, $.a_1 \cdots a_t 100 a_{t+4} \cdots = .a_1 \cdots a_t 0 b_1 b_2 \cdots$ for some b's. We can make this change in infinitely many places; each gives a different expansion. Applying this to ξ^* takes care of infinitely many 011's.

CASE III. ξ has an expansion ending in $010101 \cdots$.

Since $\xi \leqq (1+\sqrt{5})/2$ we have $.1 \leqq .(01)$. Thus $.(01) = .1 b_2 b_3 \cdots$ for some b's. We can make this change as far out in the expansion of ξ as we like, again generating infinitely many.

THEOREM 3. *For any α the set of numbers ξ having infinitely many expansions to base α is dense.*

Proof. Since $.1 < .0(b)$, there exists an n such that if $a_i = b$, $i = 1, \cdots, n$, then $.1 < .0 a_1 \cdots a_n$. Clearly any number with an expansion containing the sequence $0 a_1 \cdots a_n$ infinitely often has infinitely many expansions. But any number may be approximated arbitrarily well by such numbers.

We note that a number may even have uncountably many expansions. If $\alpha = (1+\sqrt{5})/2$, then $.1 = .011$; and the number $\xi = .(100)$ is seen to be an example of such a number.

The last result in this section deals with the relation of the algorithmic expansion of a number to its other expansions. We give the representations of a number ξ to a fixed base α the lexicographic order. Thus if $A: .a_1 a_2 \cdots$ and $B: .b_1 b_2 \cdots$ are two expansions of ξ, by $A > B$ we mean that if j is the smallest integer such that $a_j \neq b_j$, then $a_j > b_j$.

THEOREM 4. *For fixed α the expansions to base α of a number ξ form an ordered chain with greatest element. This greatest element is given by the algorithm* (3).

Proof. Since any two expansions are comparable, we have an ordered chain. Suppose $.a_1 a_2 \cdots$ is given by the algorithm and let $.b_1 b_2 \cdots$ be any other representation of ξ. We may as well assume $a_1 \neq b_1$. Then

$$a_1 = [\alpha\xi] = [b_1 + .b_2 b_3 \cdots] = b_1 + [.b_2 b_3 \cdots] \geqq b_1.$$

Thus, in fact, $a_1 > b_1$ as required.

When ξ has more than one finite expansion, however, the algorithm need

not give the shortest one. Even for rational α the algorithm may not give the finite expansion when one exists. A simple example is $\alpha = 5/2$, $\xi = .022 = .10012 \cdots$. In the next section we give a variation of the algorithm (3) which allows us to obtain effectively the periodic expansion of ξ, if one exists, when α is rational.

5. Periodic expansions to rational bases. We first state and prove necessary and sufficient conditions for a number to have a periodic expansion to a rational base. We then describe an algorithm which will indicate whether a periodic expansion exists and give it, if it does. Our proof is quite similar to that of a theorem of Kober [1; Theorem 1], which gives necessary and sufficient conditions for the algorithm (3) to yield a periodic expression, the base being rational.

If an expansion $.a_1a_2 \cdots$ is periodic, the period κ is defined as the least positive integer such that $a_{n+\kappa} = a_n$ for all sufficiently large n, and μ is defined to be the least positive integer such that $a_{n+\kappa} = a_n$ for all $n \geqq \mu$.

THEOREM 5. *Let $\alpha = P/Q$, where $(P, Q) = 1$. A number ξ has a periodic expansion to base α if and only if it is rational, say $\xi = R/S$, with $(R, S) = 1$, and is one of the numbers*

$$\sum_{j=1}^{\mu} a_j\alpha^{-j} + \alpha^{-\mu} \sum_{j=1}^{\kappa} b_j\alpha^{-j}(1 + \alpha^{-\kappa} + \alpha^{-2\kappa} + \cdots), \tag{6}$$

where $0 \leqq a_j$, $b_j \leqq [\alpha]$ and where $\mu + \kappa \leqq [SP/Q(P-Q)]$.

Proof. That the condition is sufficient is trivial. Suppose ξ has a periodic representation $.a_1a_2 \cdots$. Clearly ξ is rational; let $\xi = R/S$, $(R, S) = 1$. Define R_n for $n \geqq 0$ by

$$R_n/SP^n = \sum_{j=n+1}^{\infty} a_j\alpha^{-j} = \xi - \sum_{j=1}^{n} a_j\alpha^{-j}. \tag{7}$$

Then $0 \leqq R_n = SP^n \sum_{j=n+1}^{\infty} a_j\alpha^{-j} < SP^n\alpha^{-n}/(1-\alpha^{-1}) = SPQ^n/(P-Q)$, and

$$0 \leqq R_n/Q^n < SP/(P - Q). \tag{8}$$

Also, $R_n/P^nS + a_n/\alpha^n = R_{n-1}/P^{n-1}S$, so that

$$R_n = R_{n-1}P - a_nQ^nS \qquad \text{for } n \geqq 1. \tag{9}$$

If we suppose now that ξ has the form given by (5), we obtain

$$P^m(P^k - Q^k)R = SQ\left(\sum_{j=1}^{m} c_jQ^{j-1}P^{m-j} + Q^m \sum_{j=1}^{k} b_jQ^{j-1}P^{k-j}\right)$$

so that $Q \mid R = R_0$. With (9) this implies that $Q \mid R_1$.

Now if we let $R_1/QS = r_1/s_1$, $(r_1, s_1) = 1$, we see again that $r_1/s_1 = \alpha R_1/PS$ has a periodic expansion, namely $.a_2a_3 \cdots$, so by the above $Q \mid r_1$. Thus $Q^2 \mid R_1$.

Suppose we have proved that $Q^n \mid R_{n-1}$. By (9) we have $Q^n \mid R_n$ also. If we

define r_n and s_n by $r_n/s_n = R_n/Q^nS$, $(r_n, s_n) = 1$, then $r_n/s_n = \alpha^n R_n/P^nS = .a_{n+1}a_{n+2} \cdots$, which is periodic. Thus $Q | r_n$, and so $Q^{n+1} | R_n$. We have proved inductively that

$$(10) \qquad Q^{n+1} | R_n, \qquad n = 0, 1, \cdots .$$

Now let V_n be the nonnegative integer R_n/Q^{n+1}, let $T = [SP/Q(P-Q)]$, and consider the set $\mathcal{V} = \{V_j : j = 0, 1, \cdots, T\}$. If some element V_n of $\mathcal{V}$ is zero, then $R_n = 0$, and the expansion is finite of length no greater than n by (7). Thus $\mu + \kappa \leqq T$, as claimed. If none are zero then two are equal, since by (8) none is greater than T. Suppose that $V_p = V_q$. By (9)

$$(11) \qquad a_{n+1} = V_nP/S - V_{n+1}Q/S.$$

We note that $V_{n+1}Q/S = .a_{n+2}a_{n+3} \cdots \leqq .(b) < 2$, if $\alpha > 2$. Thus for $\alpha > 2$, $a_{n+1} = [V_nP/S]$ or $[V_nP/S] - 1$; therefore, since $V_p = V_q$,

$$(12) \qquad a_{p+1} = a_{q+1} + \epsilon,$$

where $\epsilon = 0, \pm 1$. If $\alpha < 2$, $b = 1$, and (12) is trivial. Thus (12) holds in any case; it and (11) yield

$$V_pP - V_{p+1}Q = V_qP - V_{q+1}Q + \epsilon S,$$

or $(V_{q+1} - V_{p+1})Q = \epsilon S$. But $Q | R$, so $(Q, S) = 1$; thus $Q | \epsilon$. Hence $\epsilon = 0$, $a_{p+1} = a_{q+1}$, and, by (11) $V_{p+1} = V_{q+1}$. By induction, $a_{p+n} = a_{q+n}$ for all $n \geqq 1$. Since $p, q \leqq T$, we see that the expansion is periodic and $\mu + \kappa \leqq T$.

The question of when an expansion is purely periodic, i.e. $\mu = 1$, is handled by the following corollary, which is practically a verbatim quotation of Kober's Theorem 2. If we denote R_n/Q^n by W_n, then the proof given by Kober also proves this corollary.

COROLLARY. *Let R/S have a periodic expansion to the base P/Q of the form* (6).
(*A*) *The expansion is purely periodic, i.e.* $\mu = 1$, *if and only if* $(P, S) = 1$.
(*B*) *Let $S^* = S/(P, S)$. Then κ is the smallest positive integer for which $P^\kappa \equiv Q^\kappa$ (mod S^*), and $\kappa | \phi(S^*)$, where ϕ is the Euler function.*
(*C*) *The number μ is the smallest positive integer for which $(P^{\mu-1}, S) = (P^\mu, S)$, or for which $s_{\mu-1} = S^*$.*

As our final result we obtain an algorithm for finding the periodic expansion when it exists. It is a variation of the algorithm given by (3) which relies on Lemma 2 and Theorems 1, 4, and 5 for the verification of its validity.

Let $\xi = R/S$ and $\alpha = P/Q$ as before. Set $R_0 = R$ and $b_1 = [\alpha\xi]$. The proof of Theorem 5 shows that if a periodic expansion exists, then $Q | R$. Suppose this is so. We define a_1 as b_1 or $b_1 - 1$, so as to make $R_1 = R_0P - a_1QS$ divisible by Q^2 and $0 \leqq a_1 \leqq b$, again, if this is possible. Proceeding recursively, suppose that b_{n-1}, a_{n-1}, and R_{n-1} have been defined. Then we define $b_n = [\alpha^n(\xi - .a_1 \cdots a_{n-1})]$, and

a_n as b_n or b_n-1 to make $R_n=R_{n-1}P-a_nQ^nS$ divisible by Q^{n+1}, where $0\leqq a_n\leqq b$. Since $(S, Q)=1$, it is easy to see that there is always at most one choice for a_n. Let $V_n=R_n/Q^{n+1}$.

We claim that it suffices to repeat this process $T=[SP/Q(P-Q)]$ times to determine whether ξ has a periodic expansion or not and what it is. If at any time the choice of a_n is impossible, no periodic expansion exists. If T steps may be taken, however, then a periodic expansion exists if and only if two of the V_n are equal, $n\leqq T$.

If ξ has a periodic expansion, it is easy to see that the above process produces it; and by the proof of Theorem 5 some $V_p=V_q$, where $p, q\leqq T$. On the other hand, suppose the above algorithm produces a $V_p=V_q$, $p<q$. The conditions defining a_n may be written $a_n=b_n$ or b_n+1, where $b_n=[PV_{n-1}/S]$ and $V_{n-1}P-a_nS\equiv 0 \pmod{Q}$. Since this choice is possible for a_{p+1}, and since $V_p=V_q$, the choice is clearly also possible for a_{q+1}, and $a_{p+1}=a_{q+1}$. This implies that $V_{p+1}=V_{q+1}$; by induction we see the choice is always possible and $a_{p+n}=a_{q+n}$ for all $n\geqq 1$. Thus the expansion is periodic.

The first-named author was partially supported by National Science Foundation grants 05259 and G.P. 1894; the second author was an NSF postdoctoral fellow.

The work was done while both authors were at the University of Michigan.

Reference

1. H. Kober, On quasi-decimals and on arithmetical properties of certain perfect sets and monotone functions, J. London Math. Soc., 28 (1953) 47–58.

ARITHMETIC FUNCTIONS IN AN UNUSUAL SETTING

L. CARLITZ, Duke University

1. Introduction. Let F denote an arbitrary but fixed field. By an arithmetic function is meant a mapping from the integers into F; usually the arguments are limited to positive or nonnegative integers. If f, g are two functions the sum $h=f+g$ is defined by means of

$$h(n) = f(n) + g(n). \tag{1.1}$$

There are, however, several products that are of interest. In the first place the ordinary product $h=fg$ is defined by

$$h(n) = f(n)g(n). \tag{1.2}$$

The Cauchy product is defined by

$$h(n) = \sum_{r=0}^{n} f(r)g(n-r) \tag{1.3}$$

while the Dirichlet product is defined by

$$h(n) = \sum_{rs=n} f(r)g(s). \tag{1.4}$$

For the Cauchy product n is restricted to nonnegative values while for the Dirichlet product n is restricted to positive values. For references see Bell [1].

We thus have three algebraic systems. The system based on (1.2) is a commutative ring that has zero divisors. On the other hand the systems based on (1.3) and (1.4), respectively, are domains of integrity. Moreover, the Cauchy system contains essentially a single prime, namely, the function defined by

$$f(1) = 1, \qquad f(n) = 0 \quad (n \neq 1). \tag{1.5}$$

The Dirichlet system contains infinitely many primes; it has been proved recently [2] that the unique factorization theorem holds for this system.

In the present paper we discuss a system based on another definition of product. This definition is based on the following result due to Lucas [4]. Let p be a prime and put

$$n = n_0 + n_1 p + n_2 p^2 + \cdots \qquad (0 \leqq n_j < p),$$
$$r = r_0 + r_1 p + r_2 p^2 + \cdots \qquad (0 \leqq r_j < p);$$

then

$$\binom{n}{r} \equiv \binom{n_0}{r_0}\binom{n_1}{r_1}\binom{n_2}{r_2} \cdots \pmod{p}. \tag{1.6}$$

In particular it follows from (1.6) that the binomial coefficient $\binom{n}{r}$ is prime to p if and only if

$$0 \leqq r_j \leqq n_j \qquad (j = 0, 1, 2, \cdots). \tag{1.7}$$

We accordingly define the product $h = f * g$ by means of

$$h(n) = \sum_{r=0}^{n}{}^{*} f(r)g(n-r), \tag{1.8}$$

where the summation is restricted to those r that satisfy (1.7). It may be appropriate to call h the Lucas product of f and g. The domain of definition consists of the nonnegative integers.

2. It is immediate from (1.8) that the Lucas product is both associative and commutative; also it is distributive with respect to addition as defined by (1.1). The function z defined by

$$z(n) = 0 \qquad (n = 0, 1, 2, \cdots) \tag{2.1}$$

is the identity for addition; the function u defined by

$$u(0) = 1, \qquad u(n) = 0 \quad (n > 0) \tag{2.2}$$

is the multiplicative identity for the product (1.8).

A function f is *singular* provided $f(0)=0$; otherwise it is *nonsingular*. If $f*g=u$ we call g the *inverse* of f. We now prove

THEOREM 1. *A function f possesses an inverse if and only if it is nonsingular.*

Proof. If $f*g=u$ then by (1.8) $f(0)g(0)=1$ so that $f(0)\neq 0$. Conversely if $f(0)\neq 0$, (1.8) can be thought of as a triangular system. Since the coefficient of $g(n)$ is $f(0)$, it follows that $g(n)$ is uniquely determined in terms of $g(r)$, where $0\leqq r<n$ and r satisfies (1.7).

A nonsingular function may accordingly be called a unit. Two functions that differ only by a unit factor are *associates.*

If $f*g=z$, $f\neq z$, $g\neq z$, where z is defined by (2.1), then f is a *zero divisor.*

In order to discuss zero divisors it will be convenient to define certain functions that we shall call *monomial* functions. We first define ϕ_j by means of

$$\phi_j(n)=\begin{cases}1 & (n=p^j)\\ 0 & (n\neq p^j),\end{cases} \tag{2.3}$$

where $j=0, 1, 2, \cdots$. If $m=a_0+a_1p+\cdots+a_rp^r$, $(0\leqq a_j<p)$, we define the general monomial function

$$\psi_m=\overset{a_0}{\phi_0}*\overset{a_1}{\phi_1}*\cdots*\overset{a_r}{\phi_r} \qquad (m=1, 2, 3, \cdots). \tag{2.4}$$

It is understood that f^k means $f*\cdots*f$, where the number of factors is k. Thus ψ_m is the characteristic function of $\{m\}$.

Now it is evident from (1.8) and (2.3) that

$$\overset{a}{\phi_j}(n)=\begin{cases}1 & (n=ap^j)\\ 0 & (n\neq ap^j),\end{cases} \tag{2.5}$$

where $0\leqq a<p$, while

$$\overset{p}{\phi_j}=z \qquad (j=0, 1, 2, \cdots). \tag{2.6}$$

Thus by (2.4) we have

$$\overset{p}{\psi_m}=z \qquad (m=1, 2, 3, \cdots), \tag{2.7}$$

so that every monomial function is nilpotent.

It also follows from (2.4) and (2.5) that

$$\psi_{m_1}*\psi_{m_2}=\psi_{m_1+m_2} \quad \text{or} \quad z \tag{2.8}$$

according as the binomial coefficient $\binom{m_1+m_2}{m_1}$ is or is not prime to p. More generally

$$\psi_{m_1}*\cdots*\psi_{m_k}=\psi_{m_1+\cdots+m_k} \quad \text{or} \quad z \tag{2.9}$$

according as the multinomial coefficient

$$\frac{(m_1 + m_2 + \cdots + m_k)!}{m_1!m_2! \cdots m_k!}$$

is or is not prime to p. Clearly (2.7) is a special case of (2.9).

Consider the function

$$f = f_1 * \psi_{m_1} + \cdots + f_k * \psi_{m_k}, \tag{2.10}$$

where the f_j are arbitrary functions. Raising both sides of (2.8) to the c-th power we have by the multinomial theorem

$$f^c = \sum_{r_1+\cdots+r_k=c} \frac{c!}{r_1! \cdots r_k!} f_1^{r_1} * \psi_{m_1}^{r_1} * \cdots * f_k^{r_k} * \psi_{m_k}^{r_k}.$$

For $c>k(p-1)$ at least one $r_j \geqq p$ and therefore by (2.7) we have $f^c = z$. This proves

THEOREM 2. *If $\psi_{m_1}, \cdots, \psi_{m_k}$ are monomial functions and $f_1, \cdots, f_k$ are arbitrary functions then f, as defined by* (2.10), *is nilpotent.*

We remark that an arbitrary function can be exhibited as a series of monomial functions in the following way:

$$f = f(0)u + \sum_{m=1}^{\infty} f(m)\psi_m. \tag{2.11}$$

There is of course no question of convergence; for a particular value of the argument n only one term on the right survives. We note also that if

$$g = g(0)u + \sum_{m=1}^{\infty} g(m)\psi_m, \tag{2.12}$$

then formal multiplication of (2.11) and (2.12) yields

$$f * g = f(0)g(0)u + \sum_{m=1}^{\infty} h(m)\psi_m,$$

where by (2.8)

$$h(m) = f(0)g(m) + f(m)g(0) + \sum_{r=1}^{m-1}{}^{*} f(r)g(m-r).$$

This is evidently in accord with (1.8).

3. The characterization of zero divisors apparently depends on the characteristic of the field F. We shall prove the following theorem for F of positive characteristic.

THEOREM 3. *Let F be of positive characteristic q. Then f is a zero divisor if and only if it is singular. Moreover every zero divisor is nilpotent.*

Proof. If f is a zero divisor it is evident from Theorem 1 that f must be singular. We therefore assume in what follows that $f(0)=0$. Then (2.11) becomes

$$f = \sum_{m=1}^{\infty} f(m)\psi_m. \tag{3.1}$$

It follows from the remark at the end of Section 2 and the fact that F is of characteristic q that $f^q = \sum_{m=1}^{\infty} (f(m))^q \psi_m^q$. Repeated application of this formula gives

$$f^{q^r} = \sum_{m=1}^{\infty} (f(m))^{q^r} \psi_m^{q^r} \qquad (r = 1, 2, 3, \cdots). \tag{3.2}$$

Now choose r so that $q^r \geqq p$. Then, in view of (2.7), (3.2) reduces to $f^{q^r} = z$. This evidently completes the proof of the theorem.

Theorem 3 is certainly not valid for F of characteristic 0. We can show by exhibiting an example that a singular function need not be a zero divisor. Consider the function

$$f = \sum_{m=1}^{\infty} \psi_m; \tag{3.3}$$

alternatively this can be defined by

$$f(n) = \begin{cases} 0 & (n = 0) \\ 1 & (n = 1, 2, 3, \cdots). \end{cases} \tag{3.4}$$

Assume that there exists a function $g \neq z$ such that

$$f * g = z. \tag{3.5}$$

Then g is singular and we write

$$g = \sum_{m=1}^{\infty} g(m)\psi_m. \tag{3.6}$$

It follows from (3.3) and (3.5) that

$$\sum_{r=1}^{m}{}^{*}\, g(r) = 0 \qquad (m = 1, 2, 3, \cdots), \tag{3.7}$$

where the summation is restricted to r such that $\binom{m}{r}$ is prime to p. To simplify the discussion we take $p=2$. For brevity we put

$$(j_1 j_2 \cdots j_k) = g(2^{j_1} + 2^{j_2} + \cdots + 2^{j_k}),$$

where the j's are distinct; also we shall use numerals in place of the j's. Thus, to begin with, (3.7) implies $(0)+(1)=(0)+(2)=(1)+(2)=0$, so that $g(2^j)=0$ for all j. Next, when $k=3$, we get

$$(01) + (02) + (12) = 0$$
$$(01) + (03) + (13) = 0$$
$$(02) + (03) + (23) = 0.$$

Adding these equations we get $(01)+(02)+(03)=0$. Hence we have also $(01)+(02)+(04)=0$ so that $(03)=(04)$. It follows at once that $(ij)=0$ for all i, j.

When $k=4$, (3.7) gives

$$(012) + (013) + (023) + (123) = 0$$
$$(013) + (014) + (034) + (134) = 0$$
$$(012) + (014) + (024) + (124) = 0$$
$$(023) + (024) + (034) + (234) = 0.$$

Adding we get

$$[01234] \equiv (012) + (013) + (014) + (023) + (024) + (034) = 0.$$

Similarly $[01235]=[01245]=[01345]=0$. Adding these four equations and using $[2345]=0$ we get

$$(012) + (013) + (014) + (015) = 0.$$

Then as before $(015)=(016)$ and therefore $(ijk)=0$ for all i, j, k.

It is now clear how to complete the proof that $g=z$. If we put

$$\phi = \sum_{j=0}^{\infty} \phi_j \tag{3.8}$$

then it is easily verified, when $p=2$, that

$$\begin{aligned} \phi^2 &= 2! \sum_{r<s} \phi_r\phi_s, \\ \phi^3 &= 3! \sum_{r<s<t} \phi_r\phi_s\phi_t, \cdots . \end{aligned} \tag{3.9}$$

It follows that f as defined by (3.3) satisfies

$$f = \sum_{j=1}^{\infty} \frac{1}{j!} \phi^j = \exp \phi - u.$$

Thus f and ϕ are associates. Moreover, it is evident from (3.9) that ϕ is not nilpotent.

In the general case it seems plausible that, for F of characteristic 0, a function is a zero divisor if and only if it is of the form (2.10). If this conjecture is correct it would follow that a function is a zero divisor if and only if it is nilpotent. As we have seen in Theorem 3 this is indeed the case for F of positive characteristic.

4. In the remainder of the paper we discuss briefly the L-analogs of certain well-known results. Let I be the function defined by means of

$$I(n) = 1 \qquad (n = 0, 1, 2, \cdots). \tag{4.1}$$

Since $I(0) \neq 0$ it follows from Theorem 1 that I possesses an inverse μ:

$$I * \mu = u. \tag{4.2}$$

Indeed if

$$n = a_0 + a_1 p + \cdots + a_r p^r \qquad (0 \leqq a_j < p), \tag{4.3}$$

it is easily verified that

$$\text{(4.4)} \quad \mu(n) = \prod_{j=0}^{r} \mu(a_j p^j) \qquad \text{and} \qquad \text{(4.5)} \quad \mu(ap^r) = \begin{cases} 1 & (a = 0) \\ -1 & (a = 1) \\ 0 & (a > 1), \end{cases}$$

where $0 \leqq a < p$.

We can now state the following

THEOREM 4. *If*

$$\text{(4.6)} \quad g(n) = \sum_{r=0}^{n}{}^{*} f(r) \qquad \textit{then} \qquad \text{(4.7)} \quad f(n) = \sum_{r=0}^{n}{}^{*} \mu(r) f(n-r)$$

and conversely.

The proof is an immediate consequence of (4.2). We define

$$d(n) = \sum_{r=0}^{n}{}^{*} 1, \tag{4.8}$$

so that

$$d = I^2 = I * I. \tag{4.9}$$

Then it is easily verified, with the notation (4.3), that

$$d(n) = \prod_{j=0}^{r} (a_j + 1). \tag{4.10}$$

More generally we may define

$$d_k = I^k = I * \cdots * I, \tag{4.11}$$

so that $d = d_2$ and

$$d_k(n) = \sum_{r=0}^{n} {}^{*} d_{k-1}(n) \qquad (k = 1, 2, 3, \cdots). \tag{4.12}$$

We find that

$$d_k(n) = \prod_{j=0}^{r} \binom{a_j + k - 1}{k - 1}. \tag{4.13}$$

We recall that Dickson has proved [3] that the multinomial coefficient $n!/(n_1!n_2! \cdots n_k!)$, $(n=n_1+n_2+ \cdots +n_k)$, is prime to p if and only if

$$n_i = \sum_j a_{ij}p^j \qquad (0 \leqq a_{ij} < p)$$

$$n = \sum_j a_jp^j \qquad (0 \leqq a_j < p)$$

and $a_j = \sum_{i=1}^{k} a_{ij}$ $(j=0, 1, 2, \cdots)$. It follows that the number of k-nomial coefficients (with n fixed) that are prime to p is equal to $d_k(n)$.

With n defined by (4.3) and $m=b_0+b_1+ \cdots +b_rp^r$ $(0 \leqq b_j < p)$ we define

$$(m, n)_p = \sum_{j=0}^{r} p^j \min(a_j, b_j), \tag{4.14}$$

$$[m, n]_p = \sum_{j=0}^{r} p^j \max(a_j, b_j), \tag{4.15}$$

so that $(m, n)_p + [m, n]_p = m + n$.

A function f is *factorable* provided $f(0)=1$ and

$$f(m + n) = f(m)f(n) \qquad ((m, n)_p = 0) \tag{4.16}$$

For example I, μ, d are instances of factorable functions. We have the following result.

THEOREM 5. *If f is factorable and*

$$g(n) = \sum_{r=0}^{n}{}^{*} f(r) \tag{4.17}$$

then g is factorable. Conversely if g is factorable then f is factorable.

The proof of Theorem 5 may be omitted. A more general result is the following.

THEOREM 6. *Let $h=f*g$. Then if any two of the functions f, g, h are factorable the third is also.*

A function f is *additive* provided

$$f(m + n) = f(m) + f(n) \qquad ((m, n)_p = 0). \tag{4.18}$$

If f is additive and g is defined by (4.17) we have

$$g(m + n) = \sum_{r=0}^{m}{}^{*} \sum_{s=0}^{n}{}^{*}(f(r) + f(s)) = g(m)d(n) + g(n)d(m),$$

and therefore

$$\frac{g(m+n)}{d(m+n)} = \frac{g(m)}{d(m)} + \frac{g(n)}{d(n)}.$$

Thus $g(n)/d(n)$ is additive. Now let $g(n)/d(n)$ be additive so that $g(m+n) = g(m)d(n)+g(n)d(m)$, $((m, n)_p = 0)$. Then

$$\begin{aligned} f(m+n) &= \sum_{t=0}^{m+n}{}^{*} \mu(t)g(m+n-t) \\ &= \sum_{r=0}^{m}{}^{*} \sum_{s=0}^{n}{}^{*} \mu(r)\mu(s)\{g(m-r)d(n-s) + g(n-s)d(m-r)\} \\ &= \sum_{r=0}^{n}{}^{*} \mu(r)g(m-r) \sum_{s=0}^{n}{}^{*} \mu(s)d(n-s) \\ &\quad + \sum_{r=0}^{m}{}^{*} \mu(r)d(m-r) \sum_{s=0}^{n}{}^{*} \mu(s)g(n-s) \\ &= f(m) + g(n). \end{aligned}$$

This proves

THEOREM 7. *Let* $g(n) = \sum_{r=0}^{m}{}^{*} f(r)$. *Then* $f(n)$ *is additive if and only if* $g(n)/d(n)$ *is additive.*

A straightforward computation leads to the following result.

THEOREM 8. *Let* f *and* g *be additive functions and put* $h = f * g$, $\bar{f} = I * f$, $\bar{g} = I * g$. *Then*

$$k(n) = \frac{h(n)}{d(n)} - \frac{\bar{f}(n)}{d(n)} \frac{\bar{g}(n)}{d(n)}$$

is also additive.

Supported in part by NSF grant GP-1593.

References

1. E. T. Bell, Euler algebra, Trans. Amer. Math. Soc., 25 (1923) 135–154.

2. E. D. Cashwell and C. J. Everett, The ring of number-theoretic functions, Pacific J. Math., 9 (1959) 975–985.

3. L. E. Dickson, Theorems on the residues of multinomial coefficients with respect to a prime modulus, Quart. J. Math., 33 (1902) 258–277.

4. E. Lucas, Sur les congruences des nombres eulériennes et des coefficients différentiels des fonctions trigonométriques, suivant un module premier, Bull. Soc. Math. France, 6 (1878) 49–54.

A TOPOLOGICAL APPROACH TO GEOMETRY

M. C. GEMIGNANI, SUNY at Buffalo

The purpose of this paper is to summarize a method by which at least certain topological spaces can be characterized up to homeomorphism class. The method employs the notion of a *geometry* defined on a set by means of distinguished subsets called k-flats which are generalizations of k-dimensional subspaces. The set theoretic axioms defining a geometry are intended to embody the essence of anything that mathematicians have ever called a geometry. If the set on which a geometry has been defined is also a topological space, then the topology and geometry may be axiomatically related in such a manner that the topology can be completely characterized.

We shall concern ourselves in this paper primarily with characterizing R^m as a topological space using the properties generalized from the Euclidean geometry on R^m. In R^m, the usual Euclidean geometry, the standard metric, and the algebraic structure of R^m as a vector space are tightly bound together. In our initial abstraction we do away with coordinatization, metrics, and the algebraic structure of geometries derived from linear manifolds.

A *geometry* is defined as follows: Let X be a set. An element of X is called a *point*. We define $G=\{F^{-1}, F^0, \cdots, F^n\}$ to be a geometry on X if the following axioms are satisfied:

1) F^i consists of subsets of X, $-1 \leqq i \leqq n$. An element of F^i is called an i-flat, or merely a flat.

2) The only -1-flat is the empty set, i.e. $F^{-1}=\{\varnothing\}$.

3) F^i consists of *proper* subsets of X, $-1 \leqq i \leqq n$.

4) Every set of $i+1$ points not all contained in some k-flat, $k<i$, is contained in a unique i-flat, $-1 \leqq i \leqq n$.

5) The intersection of any two flats is again a flat.

6) If $k<i$, then no i-flat is contained in any k-flat.

n is called the *length* of G and we write $l(G)=n$. i is called the *dimension* of F^i, as well as the dimension of any flat f in F^i; we write $\dim f=i$. By G^* we denote $\{F^{-1}, \cdots, F^n, F'^{n+1}\}$, where $F'^{n+1}=\{X\}$. X is then considered to be an $n+1$-flat.

The axioms defining a geometry are independent. The most trivial example of a geometry on a set X is where the i-flats of X are subsets of X of cardinality $i+1$. Clearly the usual Euclidean geometry on R^m is a geometry in our sense of length $m-1$. The great circle geometry on the 2-sphere S^2 is also a geometry on the set of points of S^2. If we let $F^{-1}=\{\varnothing\}$, $F^0=\{\{x, y\} \mid x \text{ is antipodal to } y\}$, and $F^1=\{C \mid C \text{ is a great circle}\}$, then $G=\{F^{-1}, F^0, F^1\}$ is a geometry of length 1.

The length of any geometry is assumed to be finite throughout this paper, although this is not required for many results.

Suppose now, and for the remainder of this paper, that X is a topological space on which a geometry of length $m-1 \geqq 0$ has been defined. We need a concept which will act as a link between geometry and topology, and for this purpose we use a generalized notion of convexity.

A subset W of X is said to be *convex with respect to* G if the intersection of W with any flat of G^* is connected; it follows that any convex set is connected.

We now seek an axiomatic combination of the topology and geometry on X which will force X to "approximate" R^m. The particular combination developed here gives what is called an *m-arrangement*. As the defining axioms are introduced, they will be assumed to hold in the sequel. All axioms introduced can easily be seen to hold in R^m. The first axiom in the definition of an m-arrangement is

A1: *Each* 0*-flat consists of precisely one point, i.e.* $F^0 = \{\{x\} \mid x \in X\}$.

G is said to be *topological* if every flat is closed, and the intersection of any nonempty family of convex subsets of X is also convex. We assume

A2: *G is topological.*

For any subset S of X, let $K(S)$ be the set of all convex sets which contain S and define $C(S)$, called the *convex hull* of S naturally enough, to be the intersection of all sets in $K(S)$. Since G is topological, the convex hull of S is convex and contains S provided $K(S)$ is nonempty. The following axiom is equivalent to assuming that X is convex, hence it assures us that every subset of X is contained in its convex hull.

A3: *Every* 1*-flat in* G^* *is connected.*

If $x \neq y$, then the convex hull of $\{x, y\}$ is written $\overline{xy}$ and is called the *segment* joining x and y. A subset W of X is convex iff given any two distinct points x and y of W, $\overline{xy} \subseteq W$, exactly the criterion in R^m.

If $T \subseteq X$, let $g(T)$ be the minimal flat in G^* which contains T. $g(T)$ is in fact unique. If $\dim g(T) = q$, we may write $g_q(T)$ instead of $g(T)$.

The assumptions made so far are not sufficient to prevent "pathological" 1-flats from appearing in our geometry. We therefore define X to be *locally convex* if every point of X has an open neighborhood basis consisting of sets convex with respect to G. Our fourth assumption

A4: *X is locally convex,*

assures us that every 1-flat as well as X itself is locally connected.

The following assumption prevents 1-flats from "branching":

A5: *If x, y and z are distinct points of any* 1*-flat, then* $\overline{xy} \cup \overline{yz} = \overline{xy}$, $\overline{yz}$, *and/or* $\overline{xz}$.

1-flats now look very much like connected subsets of the real line. If the topology on a 1-flat has a countable base, then that 1-flat is homeomorphic to a closed, half-open, or open interval of the real line, depending on whether it has two, one, or no noncut points, respectively.

So far the axioms have been directed at making 1-flats and segments behave. We extend our study to X as a whole using a generalized notion of simplex.

A set S of $k+1$ points is said to be *linearly independent* if S is not a subset of any i-flat, $i<k$. If $S=\{x_0, \cdots, x_k\}$ is linearly independent, then $C(S)$, the convex hull of S, is called a *k-simplex*. $C(S-\{x_i\})$ is called the ith face of $C(S)$. We set $\mathrm{d}C(S)$ equal to the union of all the faces of $C(S)$ and $\mathrm{int}C(S)=C(S)-\mathrm{d}C(S)$. The following four axioms complete the definition of an m-arrangement:

A6: *If $S=\{x_0, \cdots, x_k\}$ is linearly independent and $k\geqq 1$, then $C(S)$ is the union of all segments joining a point in the* 0*th face of $C(S)$ with x_0.*

A7: *If $C(S)$ is any k-simplex and f is any* 1*-flat in $g_k(S)$ such that f intersects the interior of any face of $C(S)$ in a single point, then f contains at least one point in* int $C(S)$.

A8: *If $C(S)$ is a k-simplex, then* Fr $C(S)$ *in* $g_k(S)\subseteq \mathrm{d}C(S)$.

A9: *If f and f' are flats with nonempty intersection, then* $\dim f+\dim f'=\dim(f\cup f')+\dim(f\cap f')$.

While this last axiom may appear a bit unjustified at first glance, it is easily seen to be true not only for geometries derived from linear manifolds, but also in the many geometries whose lattice of flats is modular. Nevertheless, for several reasons this last axiom is objectionable. The author shows in [2] that it can be replaced by two axioms which are more topological in nature.

With A1–A9, things behave in X more or less as they do in R^m. The question now is how much more we need assume so as to be able to say that X must be R^m.

We first impose a generalized parallel axiom. If f and f' are k-flats, then f is said to be *parallel* to f' if $f=f'$, or $f\cap f'=\varnothing$ and $f\vee f'$ is a $k+1$-flat. G is said to be *affine* if given any k-flat f and any point y in X, there is one and only one k-flat which contains y and is parallel to f. An m-arrangement is said to be affine if G is affine if $l(G)\geqq 1$, or if every point of X is a cut point of X if $l(G)=0$. We are now in a position to say that R^m is the only topological space which has a countable base and which admits a geometry G such that the space and G satisfy the axioms for an affine m-arrangement.

In [3] the author finds a similar characterization of the m-sphere.

This paper is based on a talk of the same title given at the October, 1964 meeting of the Indiana Section of the MAA. The material presented is part of the author's doctoral dissertation [1] at the University of Notre Dame, directed by Prof. Robert Clay. The thesis was published in its entirety in the January, 1966 issue of the *Notre Dame Journal of Formal Logic.*

References

1. M. C. Gemignani, Topological geometries and a new characterization of R^m, Notre Dame J. of Formal Logic, January, VII (1966) 57–100.

2. ———, On getting rid of an unwanted axiom in the characterization of R^m by means of topological geometries, Notre Dame J. of Formal Logic (to appear).

3. ———, A characterization of the m-sphere by means of topological geometries, Notre Dame J. of Formal Logic (to appear).

GENERAL SOLUTION OF A CLASS OF ITERATED ORDINARY LINEAR DIFFERENTIAL EQUATIONS

J. C. BURNS, Australian National University, Canberra

1. Let L be a linear differential operator defined by

$$L(y) = y^{(n)}(x) + \sum_{r=1}^{n} p_r(x)y^{(n-r)}(x), \tag{1}$$

and let the iterated linear operator L^m be defined by the relations

$$L^{s+1}(y) = L[L^s(y)], \qquad s = 1, 2, \cdots, m-1. \tag{2}$$

For a certain class of operators L, it will be shown that, provided the general solution of the equation $L(y)=0$ is known, the general solution can be found for the iterated equation

$$L^m(y) = 0. \tag{3}$$

2. The particular class of operators L for which the solution of (3) will be found is suggested by considering the m equations

$$\begin{aligned} L(y_1) &= 0, \qquad L(y_2) = y_1, \qquad L(y_3) = y_2, \cdots, \\ L(y_{m-1}) &= y_{m-2}, \qquad L(y) = y_{m-1}. \end{aligned} \tag{4}$$

If these equations can be solved, then y will be a solution of (3).

In [1], the solution of an equation of the form $L(y_2)=y_1$, where $L(y_1)=0$ is discussed, particularly when $L(y)$ is of the form

$$\begin{aligned} L(y) = y^{(n)}(x) &+ \frac{a_1}{x}y^{(n-1)}(x) + \frac{a_2}{x^2}y^{(n-2)}(x) + \cdots + \frac{a_{n-1}}{x^{n-1}}y'(x) \\ &+ \left(\frac{a_n}{x^n} - \frac{1}{\mu n}\right)y(x), \end{aligned} \tag{5}$$

where the coefficients a_i and μ are constants. For the remainder of this paper L will denote an operator of this form.

It is shown in [1] that for any function $u(x)$,

$$L(\vartheta u) = (\vartheta + n)L(u) + u/\mu, \tag{6}$$

where $(\vartheta u)(x) = xu'(x)$. It is then deduced that if $y=u$ is any solution of the equation $L(y)=0$, $y=\mu\vartheta u$ is a particular integral of the equation $L(y)=u$. It follows that both u and ϑu are solutions of the equation $L^2(y)=0$.

Let $y=u_k(x)$, $k=1, \cdots, n$, be n linearly independent solutions of the equation $L(y)=0$ so that the general solution is

$$y = \sum_{k=1}^{n} \alpha_k u_k(x). \tag{7}$$

Then the $2n$ functions $u_k, \vartheta u_k$ are solutions of the equation $L^2(y)=0$ which is of order $2n$. It is shown later that these functions are linearly independent so that the general solution of $L^2(y)=0$ is of the form

$$y = \sum_{j=0}^{1}\sum_{k=1}^{n} \alpha_{jk}\vartheta^j u_k(x) = \sum_{j=0}^{1}\sum_{k=1}^{n} \alpha_{jk} x^j u_k^{(j)}(x). \tag{8}$$

This solution of $L^2(y)=0$ has been obtained by considering the first two of equations (4). The similarity of form of the remaining equations suggests that it should be possible to generalise (7) and (8) and obtain the general solution of the equation $L^m(y)=0$ when L is of the form (5).

THEOREM. *If $L(y)$ is given by* (5) *and the general solution of $L(y)=0$ is given by* (7), *the general solution of the equation $L^m(y)=0$ is*

$$y = \sum_{j=0}^{m-1}\sum_{k=1}^{n} \alpha_{jk}\vartheta^j u_k(x) = \sum_{j=0}^{m-1}\sum_{k=1}^{n} \beta_{jk} x^j u_k^{(j)}(x), \tag{9}$$

where the arbitrary constants α_{jk}, β_{jk} can easily be related.

The proof depends on the following preliminary results.

LEMMA 1. *If $L(u)=0$, then, for $s \geqq 1$,*

$$L(\vartheta^s u) = \{(\vartheta + n)^s - \vartheta^s\} u/\mu n. \tag{10}$$

Equation (6) gives the case $s=1$ and can be used to prove (10) by induction. Let $\phi_n(\vartheta)$ denote a polynomial in ϑ of degree n with constant coefficients.

LEMMA 2. *If $L(u)=0$, then, for any $\phi_s(\vartheta)$,*

$$L[\phi_s(\vartheta)u] = \phi_{s-1}(\vartheta)u \quad \textit{for some } \phi_{s-1}(\vartheta). \tag{11}$$

This result follows at once from Lemma 1.

LEMMA 3. *If $L(u)=0$, then, for $s \geqq r+1$,*

$$L^r(\vartheta^s u) = s(s-1)\cdots(s-r+1)\mu^{-r}\vartheta^{s-r}u + \phi_{s-r-1}(u) \tag{12}$$

for some $\phi_{s-r-1}(\vartheta)$.

The case $r=1$ is simply a restatement of Lemma 1. The general result is obtained by induction making use of the case $r=1$ and of Lemma 2.

LEMMA 4. *If $L(u)=0$, then, for $s \geqq 1$,*

$$L^s(\vartheta^s u) = s!\mu^{-s}u. \tag{13}$$

Equation (12) with $r=s-1$ gives

$$L^{s-1}(\vartheta^s u) = s(s-1)\cdots 3.2\mu^{-(s-1)}\vartheta u + \phi_0 u,$$

where ϕ_0 is a constant. Since $L(u)=0$ and $L(\vartheta u)=u/\mu$, operating with L on both sides of this equation gives equation (13).

LEMMA 5. *If* $L(u)=0$ *and* $s\geqq 0$, $t\geqq 1$,

$$L^{s+t}(\vartheta^s u) = 0. \tag{14}$$

This comes at once from Lemma 4.

LEMMA 6. *If* $y=u_k(x)$ $(k=1, 2, \cdots, n)$ *are linearly independent solutions of* $L(y)=0$, *then the mn functions* $\vartheta^j u_k(x)$ $(j=0, 1, \cdots, (m-1); k=1, 2, \cdots, n)$ *are linearly independent.*

Suppose, on the contrary, that there exist constants σ_{jk}, not all zero, such that

$$\sum_{j=0}^{m-1}\sum_{k=1}^{n} \sigma_{jk}\vartheta^j u_k(x) = 0. \tag{15}$$

Operate on (15) with the operator L^{m-1}. Then, using (13) and (14), we get

$$(m-1)!\mu^{-(m-1)}\sum_{k=1}^{n} \sigma_{m-1,k}u_k(x) = 0.$$

Since the functions $u_k(x)$ are linearly independent it follows that the coefficients $\sigma_{m-1,k}$ $(k=1, 2, \cdots, n)$ all vanish. If the process is now repeated with the operators L^{m-2}, L^{m-3}, $\cdots$ in turn, it will be seen that all the coefficients σ_{jk} vanish. Since this contradicts the original assumption, the functions $\vartheta^j u_k(x)$ are linearly independent.

The proof of the theorem is now brief. For $k=1, 2, \cdots, n$, $L(u_k)=0$ and so

$$L^m(u_k) = 0 \qquad (k = 1, 2, \cdots, n). \tag{16}$$

Equation (14) with $u=u_k$ $(k=1, 2, \cdots, n)$ and $s=j$, $t=m-j$ for $j=1, 2, \cdots, m-1$ shows that

$$L^m(\vartheta^j u_k) = 0 \qquad (j = 1, 2, \cdots, m-1; k = 1, 2, \cdots, n). \tag{17}$$

(16) and (17) show that the mn functions $\vartheta^j u_k$ are solutions of the linear equation $L^m(y)=0$ which is of order mn. Lemma 6 shows that these functions are linearly independent so the general solution of the equation is an arbitrary linear combination of them.

The alternative expression for the general solution follows at once from the well-known relation

$$x^n y^{(n)} = \vartheta(\vartheta-1)(\vartheta-2)\cdots(\vartheta-n+1)y. \tag{18}$$

4. The application of the theorem can be illustrated by solving a linear equation which arises in finding periodic solutions of the iterated equation of generalised axi-symmetric potential theory (see [2], for example). In this theory

the potential $\phi(x, y)$ satisfies an equation of the form $L_k(\phi)=0$, where $L_k(\phi)$ is defined by

$$L_k(\phi) = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{k}{y}\frac{\partial\phi}{\partial y} \tag{19}$$

with x, y the axial and radial coordinates respectively. The corresponding iterated equation is $L_k^m(\phi)=0$ (see [3]).

Solutions of this iterated equation are obtained in the form

$$\phi(x, y) = \psi(y)\cos\lambda x \tag{20}$$

if $\psi(y)$ satisfies the iterated ordinary linear differential equation

$$\overset{m}{L}_{k\lambda}(\psi) = 0, \tag{21}$$

where $L_{k\lambda}$ is defined by

$$L_{k\lambda}(\psi) = \frac{d^2\psi}{dy^2} + \frac{k}{y}\frac{d\psi}{dy} - \lambda^2\psi. \tag{22}$$

$L_{k\lambda}$ is an operator of the form (5) so the theorem can be applied. Two independent solutions of $L_{k\lambda}(\psi)=0$ are

$$\psi = y^pI_p(\lambda y) \quad\text{and}\quad \psi = y^pK_p(\lambda y), \tag{23}$$

where $p=\frac{1}{2}(1-k)$ and $I_p(\lambda y)$ and $K_p(\lambda y)$ are modified Bessel functions of the first and second kind of order p with argument λy. Application of the theorem, followed by some rearrangement of the solution using relation (18) and the relation

$$\vartheta^s(x^nu) = x^n(\vartheta + n)^su \quad\text{for}\quad s \geqq 0, \tag{24}$$

gives the general solution of equation (21) as

$$\psi = y^p\sum_{t=0}^{m-1} y^t\frac{d^t}{dy^t}\{A_tI_p(\lambda y) + B_tK_p(\lambda y)\} \tag{25}$$

for arbitrary A_t, B_t.

References

1. J. C. Burns, Particular integrals for linear differential equations, this MONTHLY, 73 (1966) Slaught Paper No. 11, pp. 98–103.

2. A. Weinstein, Generalised axially symmetric potential theory, Bull. Am. Math. Soc., 59 (1953) 20–38.

3. ———, On a class of partial differential equations of even order, Ann. Mat. Pura. Appl. (Ser. 4), 39 (1955) 245–254.

THE SEMINATURAL NUMBERS

SAMUEL T. STERN, State University College at Buffalo

From the Peano axioms we can define the set of positive integers and operations of addition and multiplication thereon. From these axioms the properties of the positive integers with which we are familiar can be proved [3]. In order to admit the operation of subtraction, the set of positive integers is imbedded in a larger system, whose elements are pairs of positive integers [2].

In this paper we shall begin with four axioms which define the seminatural numbers. The fourth axiom is the principle of transfinite induction [1]. We proceed to define operations of addition and multiplication on the set of seminaturals and deduce various properties of the seminaturals from the axioms. The nonnegative integers will then represent a special case of the seminaturals. Addition and multiplication for the seminaturals will be associative and distributive but not necessarily commutative. Finally, the seminaturals will be imbedded in a larger system, whose elements are pairs of seminaturals, in which a certain kind of subtraction is possible.

DEFINITION. *Let N be a set and $<$ a binary relation on N. N is said to be a seminatural system if the following axioms are satisfied:*

AXIOM 1. *N is simply ordered with respect to $<$.*

AXIOM 2. *N is not the empty set.*

AXIOM 3. *If $x \in N$, there exists $y \in N$ such that $x<y$.*

AXIOM 4. *If G is a subset of N such that for every $x \in N$, $I(x) \subset G$ implies $x \in G$, then $G=N$. By $I(x)$ we mean $\{y \mid y \in N$ and $y<x\}$.*

Axiom 1 implies that for $x, y \in N$ exactly one of the following is true: $x=y$; $y<x$; $x<y$.

The consistency of the axioms is established by considering the nonnegative integers in their natural order, and also by considering the set

$$P = \{0, 1, 2, \cdots, 0', 1', 2', \cdots,\}$$

whose elements are ordered as they appear, i.e. $x<y$ if and only if x appears to the left of y in the above array. P is a well-ordered set, i.e. a simply ordered set such that every nonempty subset thereof has a first element. In fact, every seminatural system is well-ordered with respect to $<$. For if H is a subset of N without a first element, then for every $p \in H$ there exists $q \in H$, where $q<p$. Let $x \in N$ such that $I(x) \subset N-H$. Then $x \in N-H$, for if $x \in H$ then there exists $y \in H$ where $y<x$, and $y \notin H$. By Axiom 4, $N-H=N$ and since N is nonempty, H must be.

Because N is well-ordered and nonempty we have the following theorem.

(1) THEOREM. *Every seminatural system has a first element.*

If $x, y \in N$ such that $x<y$ and such that $x<z<y$ for no $z \in N$, we call y an

immediate successor of x and x an *immediate predecessor* of y. Axiom 1 guarantees that every element x of N has at most one immediate successor and at most one immediate predecessor. Furthermore, x has at least one immediate successor. For let $N' = \{y \mid y \in N \text{ and } x < y\}$. By Axiom 3, N' is not empty. Since N is well-ordered, N' has a first element a, which is an immediate successor of x. We denote the unique immediate successor of x by Sx.

If an element p of N has no immediate predecessor, we say p is *primary*. (1) then assures us that N contains at least one primary element. In the seminatural system P above, 0 and $0'$ are primary.

We shall call the elements of N the seminaturals. The following theorems then hold for seminaturals x, y, z and primary p.

(2) THEOREM. (i) *If $x < y$ then $Sx < Sy$.*

(ii) *If $Sx = Sy$ then $x = y$.*
(iii) *If $Sx < Sy$ then $x < y$.*
(iv) *If $x < p$ then $Sx < p$.*

Proof of (i). $x < Sx$ and $y < Sy$. If $y < Sx$ then $x < y < Sx$, which is impossible. If $y = Sx$ then $Sx < Sy$, while if $Sx < y$, then $Sx < y < Sy$, and $Sx < Sy$.

The binary operation $+$, called *addition* is defined on N by the following equations and induction (Axiom 4).

$$x + p = x, \qquad x + Sy = S(x + y). \tag{3}$$

In the system P, addition may be demonstrated as follows:

$$\begin{aligned}
2 + 0' &= 2; & 2 + 1' &= 2 + S0' = S(2 + 0') = 3\\
2' + 0 &= 2'; & 2' + 1 &= 2' + S0 = S(2' + 0) = 3'\\
2 + 0 &= 2; & 2 + 1 &= 2 + S0 = S(2 + 0) = 3\\
2' + 0' &= 2'; & 2' + 1' &= 2' + S0' = S(2' + 0') = 3'.
\end{aligned}$$

Addition in P is clearly not commutative, for $2 + 2' = 4$ while $2' + 2 = 4'$.

(4) THEOREM. $x + Sy = Sx + y$.

Proof. Let G be the set of $y \in N$ such that $x + Sy = Sx + y$ for all $x \in N$. Let $q \in N$ such that $I(q) \subset G$. If q is primary then $x + Sq = S(x + q) = Sx = Sx + q$. If $q = Sq^*$ then since $q^* \in I(q)$, $x + Sq = S(x + q) = S(x + Sq^*) = S(Sx + q^*) = Sx + Sq^* = Sx + q$. By Axiom 4, $G = N$.

(5) THEOREM. $(x + y) + z = x + (y + z)$ (associativity of addition).

Proof. Induction on z, using Axiom 4.

(6) THEOREM. $x + y + z = x + z + y$.

The last theorem implies that the left cancellation law "$a + b = a + c$ implies $b = c$" need not hold in N, for if it did addition would have to be commutative.

However, the right cancellation law holds. It is an immediate consequence of the following lemma.

(7) LEMMA. *If* $x<y$ *then* $x+z<y+z$.

Proof. Let G be the set of $z\in N$ such that $x+z<y+z$. Let $q\in N$ such that $I(q)\subset G$. If q is primary then $x+q=x$, $y+q=y$; hence $x+q<y+q$. If $q=Sq^*$ then $x+q=S(x+q^*)$ and $y+q=S(y+q^*)$. But $x+q^*<y+q^*$; hence $S(x+q^*)<S(y+q^*)$, and $x+q<y+q$.

The next theorem is an immediate consequence of (7).

(8) THEOREM. (i) *If* $x+z=y+z$ *then* $x=y$.
(ii) *If* $x+z<y+z$ *then* $x<y$.

(9) THEOREM. *If* $x<p$ *then* $x+y<p$ *for any* y.

(10) THEOREM. *If* z *is not primary then* $y<y+z$.

Proof. Let $G=A\cup B$, where A is the set of all primary elements of N and B is the set of nonprimary z of N such that $y<y+z$ for every $y\in N$. Let $q\in N$ such that $I(q)\subset G$. If q is primary then $q\in G$. If $q=Sq^*$ then if q^* is primary, $y=y+q^*<S(y+q^*)=y+q$. If q^* is not primary, $q^*\in B$ and $y<y+q^*<S(y+q^*)=y+q$. Hence $q\in G$.

The binary operation of *multiplication* is defined on N by the following equations and induction.

$$xp = p \tag{11}$$
$$xSy = xy + x.$$

Multiplication in the system P is demonstrated as follows:

$$\begin{array}{llllll}
2\cdot 0' = 0'; & 2\cdot 1' = 2\cdot S0' & = 2\cdot 0' + 2 & = 0' + 2 & = 2' \\
2'\cdot 0 = 0; & 2'\cdot 1 = 2'\cdot S0 & = 2'\cdot 0 + 2' & = 0 + 2' & = 2 \\
2\cdot 0 = 0; & 2\cdot 1 = 2\cdot S0 & = 2\cdot 0 + 2 & = 0 + 2 & = 2 \\
2'\cdot 0' = 0'; & 2'\cdot 1' = 2'\cdot S0' & = 2'\cdot 0' + 2' & = 0' + 2' & = 2'
\end{array}$$

Multiplication, like addition is not commutative in P, for $2\cdot 1'=2'$ while $1'\cdot 2=2$.

(12) THEOREM. (i) $x(y+z)=xy+xz$ (*left distributive law*).
(ii) $(x+y)z=xz+yz$ (*right distributive law*).

(13) THEOREM. $x(yz)=(xy)z$ (*associativity of multiplication*).

(14) THEOREM. *For every* x, px *is primary*.

(15) THEOREM. $(Sp)x=x$.

(16) THEOREM. $(Sx)y=xy+y$.

Proof. $(Sx)y=(S(x+p))y=(x+Sp)y=xy+(Sp)y=xy+y$.

(17) COROLLARY. $y=py+y$.

(18) THEOREM. (i) $x+yz=x+zy$, (ii) $xyz=yxz$.

The last theorem shows that the right cancellation law: "$ab=cb$, b not primary, implies $a=c$" does not have to hold in N. Otherwise multiplication would be necessarily commutative. However, the left cancellation law holds for multiplication. It follows directly from (21).

(19) THEOREM. *If $x<p$ then $zx<p$ for every z.*

(20) THEOREM. *If z is not primary then $zx<zSx$.*

(21) LEMMA. *If z is not primary and $x<y$ then $zx<zy$.*

Proof. Let z be nonprimary and G the set of y such that $x<y$ implies $zx<zy$, for all $x\in N$. Let $q\in N$, $I(q)\subset G$, and $x<q$. If q is primary then $zx<q=zq$. If $q=Sq^*$ and $x=q^*$ then $zx=zq^*<zSq^*=zq$. If however $x\neq q^*$ then either $q^*<x$ or $x<q^*$. The first case is impossible since it implies $q^*<x<q$. Thus $x<q^*$ and $zx<zq^*<zq^*+z=zq$.

(22) THEOREM. *If z is not primary and $zx=zy$, then $x=y$.*

In the theorems which follow m, n are seminaturals and p is primary.

(23) THEOREM. *If $m=p+n$ then $n=q+m$ for some primary q, namely pn.*

Proof. pn is primary by (14). But $pn+m=pn+p+n=pn+n=n$, by (17).

(24) THEOREM. *For any seminaturals m, n, either $m=x+n$ or $n=x+m$.*

Proof. Let G be the set of m such that for every $n\in N$, there exists some x such that $m=x+n$ or $n=x+m$. Let $q\in N$ and $I(q)\subset G$. If q is primary then $n=n+q$. If $q=Sq^*$ then $q^*=x'+n$ or $n=x'+q^*$ for some x'. In the first case $q=Sx'+n$. In the second case, if x' is not primary, $x'=Sx^*$ and $n=Sx^*+q^*$ $=x^*+q$; while if x' is primary, by (23), $q^*=p+n$ for some primary p, and $q=S(p+n)=Sp+n$.

(25) THEOREM. *If $a+d=c+b$ and $c+h=g+d$ then $a+h=g+b$.*

Proof. Induction on d.

(26) THEOREM. *If $a+m=a+n$ then $n=p+m$ for some primary p.*

(27) THEOREM. $pb\leqq b$.

Proof. If b is primary then $b=pb$. If b is not primary then $pb<pb+b=b$, by (10) and (17).

(28) LEMMA. *If $pb+d=pd+b$ then $b=d$.*

Proof. Suppose $b<d$. Then

$$pb+b<pd+d \qquad \text{by (17)}$$
$$pb+b+pb+d<pd+d+pd+b \qquad \text{by (7)}$$
$$pb+b+d<pd+d+b \qquad \text{by (14)}$$
$$pb<pd \qquad \text{by (6) and (8)}$$
$$pb+pb+d<pd+pd+b \qquad \text{by (7)}$$
$$pb+d<pd+b \qquad \text{by (14)},$$

contradicting the hypothesis of the lemma.

The assumption that $d<b$ leads to a contradiction in a similar manner.

(29) THEOREM. *If* $ab+xd+ad=ad+xb+ab$ *and* x *is not primary, then* $b=d$.

Proof. Induction on a.

Now we are prepared to imbed N in a system D, in which the properties of N are preserved, such that every element of D has a left additive inverse.

Let M be the set of ordered pairs (a, b) of seminaturals and define a relation "$=$" on M as follows:

$$(a, b) = (c, d) \qquad \text{if and only if } a+d=c+b. \tag{30}$$

Using the properties of N previously proven, we can show (30) to be an equivalence relation [4] on M, i.e. (30) satisfies the conditions (i) $(a, b)=(a, b)$; (ii) If $(a, b)=(c, d)$ then $(c, d)=(a, b)$; (iii) If $(a, b)=(c, d)$ and $(c, d)=(g, h)$ then $(a, b)=(g, h)$.

If we denote the equivalence class [4] containing (a, b) by $K(a, b)$ we can then define addition and multiplication on the set D, of equivalence classes of M as follows:

$$\begin{aligned} K(a, b)+K(c, d) &= K(a+c, b+d) \\ K(a, b)\cdot K(c, d) &= K(ac+bd, ad+bc). \end{aligned} \tag{31}$$

It can be shown that equations (31) uniquely define *sum* and *product* of elements of D. The associative properties of addition and multiplication and the distributive properties then follow. D contains right additive identities or *zeros*. These are elements of the form $K(m, p+m)$ or the form $K(p+n, n)$, p primary. D contains left multiplicative identities or *unities*, of the form $K(Sp, q)$ where p and q are primary. Furthermore $K(a, b)+K(x, y)=K(c, d)+K(x, y)$ implies $K(a, b)=K(c, d)$ (right cancellation law of addition) and if $K(m, n)$ is not a zero then $K(m, n)\cdot K(a, b)=K(m, n)\cdot K(c, d)$ implies $K(a, b)=K(c, d)$ (left cancellation law of multiplication).

Finally, for every element $K(a, b)$ of D there exists an element $K(a^*, b^*)$ of D such that $K(a^*, b^*)+K(a, b)$ is a zero. For if $K(a, b)$ is of the form $K(x+b, b)$

we take for $K(a^*, b^*)$ the element $K(p, x)$; while if $K(a, b)$ is of the form $K(a, x+a)$ we take $K(p+x, p)$.

It is worthwhile to note the connection between seminatural systems and the arithmetic of ordinal numbers [1]. For every seminatural system is order-isomorphic to an initial segment of the ordinals without a last element. For example, the system P is order-isomorphic to the segment of ordinals determined by $\omega+\omega$. This correspondence does not preserve the operations of addition and multiplication. Indeed, the conventional addition and multiplication for the ordinals [1] do not satisfy the right distributive law while the operations of (3) and (11) do. However, we can relate the usual operations of addition and multiplication on ordinals (>0) to the operations of (3) and (11) respectively, by means of the following equations:

$$x +_s y = x + (y - [y])$$
$$x \cdot_s y = (x - [x]) \cdot y,$$

where $+_s$ and $\cdot_s$ are the operations of (3) and (11) respectively, where $+$, $-$, $\cdot$ are the usual operations on ordinals, and where $[x]$ is the largest limit ordinal (i.e. primary) less than or equal to x, or 0 if $x<\omega$. For

$$\begin{aligned} x +_s p &= x + (p - [p]) = x + (p - p) = x + 0 = x, \\ x +_s Sy &= x + (Sy - [Sy]) = x + Sy - [y] = S(x + y - [y]) \\ &= S(x +_s y), \\ x \cdot_s p &= (x - [x])p = np = p, \qquad (0 < n < \omega), \\ x \cdot_s Sy &= (x - [x])Sy = (x - [x])(y + 1) = (x - [x])y + (x - [x]) \\ &= x \cdot_s y +_s x. \end{aligned}$$

The author wishes to thank the referee for making the observations in the final paragraph.

The author is grateful to Professor Harriet F. Montague for her assistance in the preparation of this paper.

References

1. Paul R. Halmos, Naive Set Theory, Van Nostrand, Princeton, N. J., 1960.
2. Nathan Jacobson, Lectures in Abstract Algebra, vol. 1, Van Nostrand, Princeton, N. J., 1951.
3. Edmund Landau, Foundations of Analysis, Chelsea, New York, 1957.
4. Kenneth S. Miller, Elements of Modern Abstract Algebra, Harper and Row, New York, 1958.

Mathematical Swifties

"Exactly what is Zermelo's axiom?" asked Tom, choosing his words with care.

"I'll use Schwartz' principle," Tom reflected.

"Those axioms aren't self-contradictory," said Tom consistently.

EDITED BY J. H. CURTISS, University of Miami

Address all correspondence to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457

THE SOLUTION OF A CERTAIN RECURRENCE

R. L. GRAHAM AND JOHN RIORDAN, Bell Telephone Laboratories, Murray Hill, New Jersey

In 1954, P. Turán [3] gave a proof of the identity

$$\binom{n+p}{p}^2 = \sum_{k=0}^{p} \binom{p}{k}^2 \binom{n+2p-k}{2p}$$

which he said appeared without proof in a book of the Chinese mathematician Le-Jen Shoo from 1867. This is equivalent to

$$\binom{n}{p}^2 = \sum_{k=0}^{p} \binom{p}{k}^2 \binom{n+k}{2p}$$

or

$$(1) \qquad \binom{n}{m}^2 = q_{nm} = \sum_{k=0}^{m} q_{mk} \binom{n+k}{2m}.$$

In one of the many successors to Turán's paper T. S. Nandjundiah [2] noticed that the Shoo identity is an instance of the following expansion of a product of binomial coefficients, namely

$$(2) \qquad \binom{m}{p}\binom{n}{q} = \sum_{k=0} \binom{n-m+p}{p-k}\binom{m-n+q}{k}\binom{n+k}{p+q}$$

(the upper limit of the sum is supplied by the convention that $\binom{a}{b}$ is zero if $a<0$, $b<0$, or $a<b$). Let

$$r_{nm} = \frac{1}{n+1}\binom{n-m}{m}\binom{n+1}{m+1} = \frac{1}{m+1}\binom{n}{m}\binom{n-1}{m}.$$

These numbers appeared in a study of a telephone traffic system with inputs from two sources made by John P. Runyon and are known locally as Runyon numbers; cf. J. A. Morrison [1]. It follows from (2) that

$$(m+1)r_{nm} = \binom{n}{m}\binom{n-1}{m} = \sum_{k=0} \binom{m+1}{m-k}\binom{m-1}{k}\binom{n+k}{2m}$$

or

$$(3) \qquad r_{nm} = \sum_{k=0} \frac{1}{m+1}\binom{m+1}{k+1}\binom{m-1}{k}\binom{n+k}{2m} = \sum_{k=0} r_{mk}\binom{n+k}{2m},$$

a relation similar to (1). The natural question arising is: what is the general solution of

$$x_{nm} = \sum_{k=0}^{m} x_{mk}\binom{n+k}{2m}. \tag{4}$$

Since the recurrence (4) leaves x_{nn} undetermined, this is the same as asking for the coefficient $X_k(n, m)$ in

$$x_{nm} = \sum_{k=0}^{m} X_k(n, m)x_{kk}. \tag{4a}$$

The answer is given by the following

THEOREM. *If* $n=0, 1, 2, \cdots, m=0, 1, \cdots, n$, *and*

$$x_{nm} = \sum_{k=0}^{m} x_{mk}\binom{n+k}{2m}, \tag{4}$$

then

$$x_{nm} = \sum_{k=0}^{m} \frac{2k+1}{m+k+1}\binom{n+k}{m+k}\binom{n-1-k}{m-k}x_{kk}, \qquad \text{for } m < n \tag{5}$$

with arbitrary x_{kk}.

For a proof of the theorem, notice first that when $x_{nm}=r_{nm}$, $x_{kk}=r_{kk}=\delta_{0k}$, with δ_{nm} the Kronecker delta; hence

$$X_0(n, m) = r_{nm} = \frac{1}{m+1}\binom{n}{m}\binom{n-1}{m}.$$

Next, suppose that

$$x_{nm} = \frac{2p+1}{m+p+1}\binom{n-1-p}{m-p}\binom{n+p}{m+p}, \qquad p = 1, 2, \cdots, m.$$

Then, by (2)

$$\begin{aligned} x_{nm} &= \sum_{k=0} \frac{2p+1}{m+p+1}\binom{m-1-p}{k-p}\binom{m+p+1}{k+p+1}\binom{n+k}{2m} \\ &= \sum_{k=0} \frac{2p+1}{k+p+1}\binom{m-1-p}{k-p}\binom{m+p}{k+p}\binom{n+k}{2m} \\ &= \sum_{k=0} x_{mk}\binom{n+k}{2m} \end{aligned}$$

while $x_{kk}=\delta_{pk}$; hence

$$X_p(n, m) = \frac{2p+1}{m+p+1}\binom{n-1-p}{m-p}\binom{n+p}{m+p}, \qquad p = 0, 1, \cdots, m$$

and the theorem is proved.

The theorem leads to binomial identities whenever a particular solution of (4) (for which $x_{kk} \neq \delta_{pk}$, $p=0, 1, \cdots, m$) is known. Thus in the first instance $x_{nm} = q_{nm}$ yields

$$\binom{n}{m}^2 = \sum_{k=0}^{m} \frac{2k+1}{m+k+1}\binom{n-1-k}{m-k}\binom{n+k}{m+k} = \sum_{k=0}^{m} X_k(n, m)$$

since $q_{nn}=1$.

A direct proof of this identity is as follows. First

$$\begin{aligned}
\sum_0^m \frac{2k+1}{m+k+1}&\binom{n-1-k}{m-k}\binom{n+k}{m+k} \\
&= \sum_0^m \frac{2k+1}{n-m}\binom{n-1-k}{m-k}\binom{n+k}{m+k+1} \\
&= \sum_0^m \frac{2m+1}{n-m}\binom{n-1-k}{m-k}\binom{n+k}{m+k+1} \\
&\quad - 2\sum_0^m \frac{m-k}{n-m}\binom{n-1-k}{m-k}\binom{n+k}{m+k+1} \\
&= f_{nm} - g_{nm}.
\end{aligned}$$

Next we have

$$\begin{aligned}
f_{nm} &= \frac{2m+1}{n-m}\sum_0^m \binom{n-m+k-1}{k}\binom{n+m-k}{2m+1-k} \\
&= \frac{2m+1}{n-m}\sum_{m+1}^{2m+1} \binom{n-m+k-1}{k}\binom{n+m-k}{2m+1-k} \\
&= \frac{2m+1}{2n-2m}\sum_0^{2m+1} \binom{n-m+k-1}{k}\binom{n+m-k}{2m+1-k} \\
&= \frac{2m+1}{2n-2m}\binom{2n}{2m+1} = \binom{2n}{2m}
\end{aligned}$$

(the next to last step uses one form of the Vandermonde relation). Also

$$g_{nm} = 2\sum_1^m \binom{n-1-k}{m-1-k}\binom{n+k}{m+k+1} = 2\sum_0^{m-1}\binom{n-m+k}{k}\binom{n+m-1-k}{2m-k}$$

and

$$\begin{aligned}
\binom{2n}{2m} &= \sum_0^{2m} \binom{n-m+k}{k}\binom{n+m-1-k}{2m-k} \\
&= \sum_0^{m-1} \binom{n-m+k}{k}\binom{n+m-1-k}{2m-k} \\
&\quad + \sum_0^{m} \binom{n-m-1+k}{k}\binom{n+m-k}{2m-k} \\
&= \tfrac{1}{2}\cdot q_{nm} = \sum_0^{m} \binom{n-m-1+k}{k}\left[\binom{n+m-k-1}{2m-k}\right. \\
&\quad \left. + \binom{n+m-k-2}{2m-k-1} + \cdots + \binom{n}{m+1} + \binom{n}{m}\right] \\
&= g_{nm} + \binom{n}{m}^2
\end{aligned}$$

which proves the identity.

Notice that

$$(2m+1)^{-1}f_{nm} = (2m+1)^{-1}\binom{2n}{2m} = \sum_0^{m} \frac{1}{m+k+1}\binom{n-1-k}{m-k}\binom{n+k}{m+k}$$

which is equation (5) with $x_{kk}=(2k+1)^{-1}$; hence

$$x_{nm} = (2m+1)^{-1}\binom{2n}{2m}$$

is a solution of (4) and

$$\frac{1}{2m+1}\binom{2n}{2m} = \sum_0^{m} \frac{1}{2k+1}\binom{2m}{2k}\binom{n+k}{2m}$$

or

$$\binom{2n}{2m} = \sum_0^{m} \binom{2m+1}{2k+1}\binom{n+k}{2m}.$$

A further result, which we do not take space to prove, is

$$\binom{2n+1}{2m} = \sum_0^{m} \binom{2m+1}{2k}\binom{n+k}{2m}$$

which is the x_{nm} with $x_{kk}=2k+1$. Since sums and differences of solutions of (4)

are also solutions, it follows that

$$x_{nm} = \frac{1}{2}\left[\binom{2n+1}{2m} - \binom{n}{m}^2\right]$$

is the solution for which $x_{kk} = k$.

References

1. J. A. Morrison, A certain functional-difference equation, Duke Math. J., 31 (1964) 445–448.
2. T. S. Nandjundiah, Remark on a note of P. Turán, this MONTHLY, 65 (1958) 354.
3. P. Turán, On a problem in the history of Chinese mathematics, Mat. Lapok, 5 (1954) 1–6.

ON THE TOTIENT FUNCTIONS OF JORDAN AND ZSIGMONDY

J. E. SHOCKLEY, University of Wyoming, AND R. J. HURSEY, Madison College

Introduction. K. Zsigmondy (see [2], p. 152) devised a function to determine the number of elements of a certain order in a finite abelian group.

In this note it will be shown that Zsigmondy's function can be described completely by use of Jordan's totient function (see [2], p. 147). The proof is elementary and is much simpler than the lengthy combinatorial proofs of the formula found in the literature (see, for example, [1]).

I. In order to translate the problem into number-theoretic concepts, we make the following definitions:

DEFINITION. *Let n and k be positive integers. A k-tuple $\{a_1, a_2, \cdots, a_k\}$ of positive integers is called a prime sequence for n (of length k) provided $1 \leqq a_i \leqq n$ and $(a_1, a_2, \cdots, a_k, n) = 1$ (the parentheses denote the greatest common divisor).*

DEFINITION. *If n and k are positive integers, then $J_k(n)$ denotes the number of distinct prime sequences for n, each of length k. $J_0(n)$ is defined to be zero.*

DEFINITION. *Let $m, n_1, n_2, \cdots, n_s$ be fixed positive integers. An s-tuple $\{a_1, a_2, \cdots, a_s\}$ of positive integers is called a primitive sequence for m (with respect to $n_1, \cdots, n_s$) provided*
(1) $1 \leqq a_i \leqq n_i$ $(i = 1, 2, \cdots, s)$ *and*
(2) *m is the smallest positive integer such that* $ma_i \equiv 0 \pmod{n_i}$ $(i = 1, 2, \cdots, s)$.

DEFINITION. *If m is a positive integer then $\psi(m) = \psi(m; n_1, n_2, \cdots, n_s)$ denotes the number of distinct primitive sequences for m (with respect to $n_1, n_2, \cdots, n_s$).*

Thus if G is a finite abelian group with independent generators $g_1, g_2, \cdots, g_s$ of order $n_1, n_2, \cdots, n_s$, respectively, then $\psi(m)$ is the number of elements of G of order m.

II. THEOREM. *ψ is a multiplicative function.*

Proof. Let a and b be relatively prime positive integers, let $\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ and $\{\beta_1, \beta_2, \cdots, \beta_s\}$ be primitive sequences for a and b, respectively. Let γ_i be the smallest positive solution of

$$(3) \qquad \begin{aligned} x &\equiv \alpha_i \pmod{n_i/(n_i, b)} \\ x &\equiv \beta_i \pmod{n_i/(n_i, a)}. \end{aligned}$$

Since $(a, b) = 1$, it follows from the Chinese Remainder Theorem that γ_i is the unique solution of (3) in the range $1 \leqq x \leqq \text{l.c.m. } [n_i/(n_i, a), n_i/(n_i, b)] = n_i$. If y is the smallest positive integer such that

$$(4) \qquad y\gamma_i \equiv 0 \pmod{n_i} \qquad (i = 1, 2, \cdots, s)$$

then $y \leqq ab$ since obviously $ab\gamma_i \equiv 0 \pmod{n_i}$. From (3) and (4) we see that $y\alpha_i \equiv 0 \pmod{n_i/(n_i, b)}$. From the fact that $ya\alpha_i \equiv 0 \pmod{n_i}$ we see that $y\alpha_i \equiv 0 \pmod{n_i/(n_i, a)}$. Since l.c.m. $[n_i/(n_i, a), n_i/(n_i, b)] = n_i$, it follows that $y\alpha_i \equiv 0 \pmod{n_i}$. But since $\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ is a primitive sequence for a, this implies that a divides y. Similarly, b divides y and thus $y = ab$, so that $\{\gamma_1, \gamma_2, \cdots, \gamma_s\}$ is a primitive sequence for ab.

On the other hand, if we are given that $\{\gamma_1, \gamma_2, \cdots, \gamma_s\}$ is a primitive sequence for ab and we define α_i and β_i by

$$\begin{aligned} \alpha_i &\equiv \gamma_i \pmod{n_i/(n_i, b)}, \qquad 1 \leqq \alpha_i \leqq n_i/(n_i, b) \\ \beta_i &\equiv \gamma_i \pmod{n_i/(n_i, a)}, \qquad 1 \leqq \beta_i \leqq n_i/(n_i, a) \end{aligned}$$

it follows easily that $\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ and $\{\beta_1, \beta_2, \cdots, \beta_s\}$ are primitive sequences for a and b, respectively, so that the above method yields all primitive sequences for ab. Thus $\psi(ab)$, the number of primitive sequences for ab, is equal to $\psi(a) \cdot \psi(b)$.

THEOREM. *If p is a prime, p^α divides n_i $(i = 1, 2, \cdots, k)$ and p^α does not divide n_i $(i = k+1, k+2, \cdots, s)$ then*

$$\psi(p^\alpha) = J_k(p^\alpha) \prod_{k+1}^{s} \delta_i, \qquad \textit{where } \delta_i = (p^\alpha, n_i).$$

Proof. We shall show that $\{a_1, a_2, \cdots, a_s\}$ is a primitive sequence for p^α if and only if

$$(5) \qquad \{a_1p^\alpha/n_1, a_2p^\alpha/n_2, \cdots, a_kp^\alpha/n_k\}$$

is a prime sequence for p^α of length k and

$$(6) \qquad a_i \equiv 0 \pmod{n_i/\delta_i} \qquad (i = k+1, k+2, \cdots, s).$$

The proof will be in two parts.

Part 1. Assume that $\{a_1, a_2, \cdots, a_s\}$ is a primitive sequence for p^α. It is obvious that (6) holds. If $(a_1p^\alpha/n_1, \cdots, a_kp^\alpha/n_k, p^\alpha) = p^\beta$ $(\beta \geqq 1)$, then $a_ip^{\alpha-\beta} \equiv 0 \pmod{n_i}$ $(i = 1, 2, \cdots, k)$. Since p^α does not divide n_i $(i = k+1, \cdots, s)$,

if we let $p^\lambda = \max\{p^{\alpha-\beta}, \delta_{k+1}, \cdots, \delta_s\}$, then $p^\lambda < p^\alpha$ and $a_i p^\lambda \equiv 0 \pmod{n_i}$ $(i=1, 2, \cdots, s)$ which contradicts the fact that $\{a_1, a_2, \cdots, a_s\}$ is a primitive sequence for p^α. Thus (5) must hold. (A proof similar to the above shows that if $k=0$ then no primitive sequences exist for p^α. In this case $\psi(p^\alpha) = J_0(p^\alpha) \cdot \delta_1\delta_2 \cdots \delta_s = 0$.)

Part 2. Suppose (5) and (6) hold. Then obviously p^α is a solution of the system of congruences $xa_i \equiv 0 \pmod{n_i}$ $(i=1, 2, \cdots, s)$. If t is the smallest positive solution of the system, then obviously $t \leqq p^\alpha$. Since (5) holds there exist integers $x_1, x_2, \cdots, x_{k+1}$ such that

$$x_1 a_1 p^\alpha / n_1 + \cdots + x_k a_k p^\alpha / n_k + x_{k+1} p^\alpha = 1.$$

But then

$$x_1 p^\alpha (t a_1 / n_1) + \cdots + x_1 p^\alpha (t a_1 / n_1) + p^\alpha t x_{k+1} = t,$$

which, since n_i divides ta_i, implies that p^α divides t. Thus $t = p^\alpha$ so that $\{a_1, a_2, \cdots, a_s\}$ is a primitive sequence for p^α.

Since there are $J_k(p^\alpha)$ choices for the first k terms in a primitive sequence for p^α, δ_{k+1} choices for the $k+1$-st term, *etc.*, it follows that $\psi(p^\alpha) = J_k(p^\alpha)\delta_{k+1} \cdots \delta_s$.

The two theorems together imply that Jordan's function $J_k(n)$ is a multiplicative function. The well-known fact that $J_k(p^\alpha) = p^{\alpha k}(1 - 1/p^k)$ (see [2] p. 147) completes the evaluation of Zsigmondy's function, ψ.

References

1. R. D. Carmichael, Generalizations of Euler's ϕ-Function, With Application to Abelian Groups, Quart. J. Pure and Appl. Math., 44 (1913) 94–104.

2. L. E. Dickson, History of the Theory of Numbers, Vol. I, Carnegie Institute of Washington, 1919.

A SOLUTION OF A SINGULAR, MIXED PROBLEM FOR THE EQUATION OF EULER-POISSON-DARBOUX (EPD)

B. A. Fusaro, University of South Florida

Introduction. Let $E = D_{tt} + (k/t) \cdot D_t - D_{xx}$, $0 < k$, and consider the singular, mixed problem

$$(1) \qquad E(u) = 0, \qquad 0 < x < \pi, \qquad 0 < t;$$

$$(2) \qquad u(x, 0) = f(x), \qquad f \in C'', \qquad u_t(x, 0) = 0, \qquad 0 \leqq x \leqq \pi;$$

$$(3) \qquad u(0, t) = u(\pi, t) = 0, \qquad 0 \leqq t.$$

Of course f satisfies the boundary data (3), and the value of t in the initial data (2) must be understood in the limiting sense $t \to 0+$.

Equation (1) is classical, and an exposition of its theory appears in Darboux [3]. The singular, initial value problem (1, 2) has been studied extensively in m space dimensions and for general k since A. Weinstein's work in 1952 [4].

In this note a series solution of the singular, mixed EPD problem (1, 2, 3) is found by elementary means, and the series is reduced to a closed form. This solution is shown to be unique.

The solution is then expressed as the mean value of the initial datum f, taking the form of the known solution of the initial value problem (1, 2).

A physical interpretation, and a regular, mixed problem are discussed briefly.

A formal solution. Assume there exists a solution of the form $u(x, t) = X(x)T(t)$ and substitute it in (1) to get $X''/X = (k/t)\dot{T}/T + \ddot{T}/T$. By a familiar method, a formal solution of the problem turns out to be

$$(4) \qquad u(x, t) = \Gamma(1 - p) \sum_{n=1}^{\infty} b_n \sin nx \cdot (nt/2)^p J_{-p}(nt), \qquad 2p = 1 - k,$$

where the b's are the Fourier sine coefficients of f:

$$b_n = (2/\pi) \int_0^{\pi} f(x) \sin nx dx, \quad \text{and} \quad J_{-p}(z) = \sum_{j=0}^{\infty} (-1)^j (z/2)^{2j-p}/\gamma_j,$$

$\gamma_j = \Gamma(j+1)\Gamma(j-p+1)$, the Bessel function of index $-p$.

If $k = 0$, then $u(x, t) = \Gamma(1/2) \sum_{n=1}^{\infty} b_n \sin nx (nt/2)^{1/2} J_{-1/2}(nt)$. Substituting $\Gamma(1/2) = \sqrt{\pi}$ and $J_{-1/2}(z) = \sqrt{2/(\pi z)} \cos z$ yields $u(x, t) = \sum_{n=1}^{\infty} b_n \sin nx \cos nt$, the known solution for the wave equation with data (2, 3). (See [1, 2].)

The solution in closed form. The following integral relation [1, p. 482] will be needed:

$$\Gamma(1/2)\Gamma(k/2)(z/2)^p J_{-p}(z) = 2 \int_0^{\pi/2} \cos(z \sin \tau) \cos^{k-1} \tau d\tau,$$

absolutely and uniformly convergent for $0 < k = 1 - 2p$.

The formal solution (4) is well defined since the series is absolutely and uniformly convergent for $0 < t$, $0 \leqq x \leqq \pi$. This convergence follows from an application of the above integral relation and the absolute and uniform convergence of $\sum_{n=1}^{\infty} b_n \sin nx = f(x)$, ($f \in C''$, by hypothesis).

The series representation of f will now be used to extend f as an odd function of period 2π.

Substitute the integral relation in (4) and interchange the order of summation and integration to get $u(x, t) = (2/B)\int_0^{\pi/2} \cos^{k-1} \tau \sum_{n=1}^{\infty} b_n \sin nx \cos nst \, d\tau$, where $s = \sin \tau$ and the Gamma functions have been replaced by the Beta function $B = B(1/2, k/2)$. Write $2 \sin nx \cos nst = \sin n(x + st) + \sin n(x - st)$ and distribute the sum in the previous expression for u to get

$$(5) \qquad u(x, t) = (1/B) \int_0^{\pi/2} \cos^{k-1} \tau [f(x + st) + f(x - st)] d\tau, \qquad s = \sin \tau.$$

It will be verified directly that (5) is a solution of the mixed problem (1, 2, 3).

That $u(x, 0) = f(x)$ follows from writing

$$2\int_0^{\pi/2} \cos^{k-1} \tau d\tau = 2\int_0^{\pi/2} \sin^{2\sigma-1} \tau \cos^{2\nu-1} \tau d\tau = B(\sigma, \nu),$$

where $\sigma = 1/2$ and $\nu = k/2$; and clearly $u_t(x, 0) = 0$, verifying the initial data. The boundary data are satisfied because f is odd (left point) and also periodic (right point). Finally,

$$\begin{aligned}
-tE(u) &= t(u_{xx} - u_{tt} - (k/t)u_t) \\
&= (t/B)\int_0^{\pi/2} \{[f''(x + st) + f''(x - st)] \cos^2 \tau - ks/t[f'(x + st) - f'(x - st)]\} \\
&\quad \cdot \cos^{k-1} \tau d\tau \\
&= (1/B)\int_0^{\pi/2} \{t[f''(x + t \sin \tau) + f''(x - t \sin \tau)] \cos^{k+1} \tau \\
&\quad - k[f'(x + t \sin \tau) + f'(x - t \sin \tau)] \cos^{k-1} \tau \sin \tau\} d\tau \\
&= (1/B)\int_0^{\pi/2} d/d\tau\{[f'(x + t \sin \tau) - f'(x - t \sin \tau)] \cos^k \tau\} d\tau = 0.
\end{aligned}$$

Uniqueness. An energy integral argument [2] will yield uniqueness for possible solutions u with the following properties: (i) $u \in C'$ for $0 \leqq t$; (ii) $u \in C''$ for $0 < t$; (iii) u behaves no worse than $t^{\beta+1}$, $0 < \beta$, at $t = 0_+$. The first two properties are the usual sort required in such arguments, while the third is necessary to induce convergence in an integral arising from the term peculiar to the EPD equation, k/t. The function u defined by (5) clearly satisfies properties (i), (ii), (iii).

Assume there exists another solution v and put $w = u - v$, so that w is a solution of the completely homogeneous mixed problem. It will be shown that $w = 0$.

The following expression is integrated over the rectangle $R = [0, \pi]x[\delta, T]$, $0 < \delta < T$; $2w_tE(w) = 2w_t(w_{tt} + (k/t)w_t - w_{xx}) = (w_x^2 + w_t^2)_t - 2(w_xw_t)_x + 2kw_t^2/t$. Apply Green's theorem to get $2\iint_R w_tE(w)dxdt = -\oint [(w_x^2 + w_t^2)dx + 2w_xw_tdt] + 2k\iint_R w_t^2/tdxdt$. The line integral can be written as two ordinary integrals,

$$\int_0^\pi \{[w_x^2(x, T) + w_t^2(x, T)] - [w_x^2(x, \delta) + w_t^2(x, \delta)]\} dx$$
$$+ 2\int_\delta^T [w_x(0, t)w_t(0, t) - w_x(\pi, t)w_t(\pi, t)]dt,$$

the last of which vanishes since $w_t(0, t) = w_t(\pi, t) = 0$, due to the homogeneous boundary data. Now taking the limit as $\delta \to 0$ and noting that $w_x(x, 0) = w_t(x, 0) = 0$ due to the homogeneous initial data, there results $0 = 2\int_0^T dt \int_0^\pi w_tE(w)dx = \int_0^\pi (w_x^2 + w_t^2)dx + 2k\int_0^T dt \int_0^\pi w_t^2/t\, dx$. Clearly $w_x = w_t = 0$, so that $dw = 0$, or $w =$ constant; and by the initial data, $w = 0$.

The solution as a weighted mean value of f. The solution of the initial value problem (1, 2) is

$$u(x, t) = 2/B \int_0^1 (1 - \rho^2)^{(k-2)/2} M(x, \rho t; f) d\rho, \tag{6}$$

where $M(x, \sigma; f) = [f(x+\sigma) + f(x-\sigma)]/2$ (see [4]).

This formula coincides with the solution (5), providing that f is interpreted as in the third section of this note, as can be seen after the change of variable $\rho = \sin \tau$. Beginning with (6) and a suitable interpretation of f, it could be verified that (6) is a solution of the mixed problem by simply repeating the reasoning of the third section. This observation, due to Prof. K. Stellmacher, suggested an extension to more than one space variable, and it is hoped that some results on this interesting generalization will soon be ready for publication.

A physical interpretation. Equation (1) will be replaced by

$$u_{tt} + [k/(t + \delta)]u_t - u_{xx} = 0, \tag{7}$$

yielding the regular, mixed EPD problem (1, 2, 3) which can be interpreted as the problem of a vibrating string with fixed end points and with a damping factor $k/(t+\delta)$; denote the solution of this problem by $u(x, t; \delta)$. It will be assumed that $1-2p = k < 1$ to avoid integral p and Neumann functions.

Techniques similar to those indicated above will yield $u(x, t; \delta) = \sum_{n=1}^{\infty} b_n \sin nx \cdot T_n(t; \delta)$, where

$$T_n(t; \delta) = (\tau/\delta)^p [J_{1-p} J_p(n\tau) + J_{p-1} J_{-p}(n\tau)] / [J_{1-p} J_p + J_{p-1} J_{-p}],$$

$\tau = t + \delta$, and the J's without arguments are evaluated at $n\delta$. It can be shown directly from the definition of the Bessel functions that as $\delta \to 0$, $u(x, t; \delta) \to u(x, t)$.

Thus the original singular problem (1, 2, 3) can be interpreted as a vibrating string with fixed end points and damping factor k/t, which is to say that the medium has an initially infinite viscosity.

Presented at the Annual Meeting of the Southeastern Section on March 21, 1964, at Charleston, S. C., under a similar title.

References

1. R. Courant and D. Hilbert, Methods of Mathematical Physics, vol. I, Interscience, New York, 1953.

2. ———, *ibid.* vol. II, 1962.

3. G. Darboux, Leçons sur la Théorie Générale des Surfaces, vol. II, 2nd ed., Gauthier-Villars, Paris, 1914–1915, pp. 54 ff.

4. A. Weinstein, On the wave equation and the equation of Euler-Poisson, Proc. Symp. Appl. Math., vol. V, McGraw-Hill, New York, 1954.

SOLUTIONS OF LINEAR DIFFERENTIAL SYSTEMS SATISFYING MULTIPLE-POINT BOUNDARY CONDITIONS

J. B. GARNER, Louisiana Polytechnic Institute

I. Introduction. We consider the boundary value problem

$$(1)\qquad y' = A(x)y + b(x), \qquad y = \mathrm{col}(y_i), \qquad b(x) = \mathrm{col}(b_i(x)), \qquad A(x) = (a_{ij}(x))$$

and

$$(2)\qquad y_1(\alpha_1) = \beta_1,\ y_m(\alpha_2) = \beta_m, \qquad y_n(\alpha_3) = \beta_n, \qquad m = 2, \cdots, n-1,$$

where $A(x)$ and $b(x)$ are continuous over $[a, b]$, where $\alpha_1, \alpha_2, \alpha_3$ belong to $[a, b]$, and where the β_i are arbitrarily assigned real numbers. This is the same problem considered in [1], [2], and [3].

In the present note we establish additional results on the problem. The approach is similar to that taken in [3] and, for this reason, some of the proofs are omitted. The results here imply the results of [3] for n even. Otherwise, these results neither imply nor are implied by the results of the preceding papers.

II. The Multiple-Point Problem. For points $\alpha_1, \alpha_2, \alpha_3$ of $[a, b]$ such that $\alpha_1 \leqq \alpha_2 \leqq \alpha_3$, with one equality holding if $n=2$, we define the following conditions:

A. On (α_1, α_3): $a_{ij}(x) \equiv 0$ if $i \geqq j$, except that $a_{n1} \geqq 0$; $a_{1n} \geqq 0$; $a_{ij}(x) \geqq 0$ on (α_2, α_3) and does not change sign except possibly at α_2.

B. For each fixed j, $2 \leqq j \leqq n-1$, $a_{ij}(x)a_{i,j+1}(x)a_{j,j+1}(x) \leqq 0$ for $i=1, 2, \cdots, j-1$ and $x \in (\alpha_1, \alpha_2)$. Identically vanishing products are considered for uniqueness to be nonpositive.

C. For each fixed i, $1 \leqq i \leqq n$, there exist segments $(\alpha_2-\delta_i, \alpha_2)$ and $(\alpha_2, \alpha_2+\delta_i)$ on which $a_{i1}(x), \cdots, a_{in}(x)$ do not all have a common zero.

Condition (B) is considerably less restrictive than the corresponding condition of [3] since it, with (A), allows 2^{n-2} different combinations of signs. Here the signs of any $n-2$ functions determine uniquely the signs of the remaining functions. The difference in proofs of this note and those of [3] occurs in

LEMMA 1. *Let* (A), (B), *and* (C) *hold for some* $\alpha_1, \alpha_2, \alpha_3$ *of* $[a, b]$ *and let* $Y(x)$ *be the* $n \times n$ *solution of*

$$(3)\qquad Y' = A(x)Y, \qquad Y(\alpha_2) = E\ (\textit{the } n \times n \textit{ identity matrix}).$$

Then, for $h=1$ *or* n, $x=\alpha_2$ *is an isolated zero of* $y_{mh}(x)$, $m=1, \cdots, n$; $m \neq h$.

Proof. We give the proof for $h=1$; the proof for $h=n$ follows in a similar manner. From (3), $y_{11}(\alpha_2)=1$; hence $y_{11}(x)>0$ on $(\alpha_2-\delta, \alpha_2+\delta)$ for some $\delta>0$. Thus

$$y_{n1}'(x) \equiv a_{n1}(x)y_{11}(x) > 0 \qquad \text{on } (\alpha_2 - \epsilon_n, \alpha_2),\ (\alpha_2, \alpha_2 + \epsilon_n),$$

where $\epsilon_n = \min(\delta, \delta_n)$. Since $y_{n1}(\alpha_2)=0$, $y_{n1}(x)<0$ on $(\alpha_2-\epsilon_n, \alpha_2)$, $y_{n1}(x)>0$ on

$(\alpha_2, \alpha_2+\epsilon_n)$. Since $y_{n-1,1}(\alpha_2)=0$ and $y'_{n-1,1}(x)\equiv a_{n-1,n}(x)y_{n1}(x)$, we now have, by (C),

$$a_{n-1,n}(x)y_{n-1,1}(x) > 0 \qquad \text{on } (\alpha_2 - \epsilon_{n-1}, \alpha_2), (\alpha_2, \alpha_2 + \epsilon_{n-1}),$$

where $\epsilon_{n-1}=\min(\epsilon_n, \delta_{n-1})$. Thus, by (B), $a_{k,n-1}(x)y_{n-1,1}(x)$ and $a_{kn}(x)y_{n1}(x)$ agree in sign over these segments for each fixed k, $1\leqq k\leqq n-2$. Also, for $k=n-2$, (C) implies that the sum of these terms is nonzero in $(\alpha_2-\epsilon_{n-2}, \alpha_2)$ and $(\alpha_2, \alpha_2+\epsilon_{n-2})$, where $\epsilon_{n-2}=\min(\epsilon_{n-1}, \delta_{n-2})$. Now $y'_{n-2,1}(x)$, thus $y_{n-2,1}(x)$, is nonzero in these segments.

Assume that, by continuing in this manner, we find that $y_{i1}(x)=0$ over $(\alpha_2-\epsilon_p, \alpha_2)$, $(\alpha_2, \alpha_2+\epsilon_p)$ and that

$$a_{i,p+1}(x)y_{p+1,1}(x), \cdots, a_{in}(x)y_{n1}(x) \tag{4}$$

agree in sign for each fixed i, $1\leqq i\leqq p$, where $p\geqq 3$. The equation from (1) for $y'_{p1}(x)$, together with (B) and (3), imply that $a_{ip}(x)y_{p1}(x)$ and $a_{i,p+1}(x)y_{p+1,1}(x)$ agree in sign. Hence, this with (C), (3), and (4) imply that $y'_{p-1,1}(x)$ is nonzero over $(\alpha_2-\epsilon_{p-1}, \alpha_2)$, $(\alpha_2, \alpha_2+\epsilon_{p-1})$, where $\epsilon_{p-1}=\min(\epsilon_p, \delta_{p-1})$. This implies that $y_{p-1,1}(x)\neq 0$ over these segments since $y_{p-1,1}(\alpha_2)=0$.

Thus each $y_{m1}(x)$, $m=2, \cdots, n$, is either positive or negative to the immediate left and right of $x=\alpha_2$, and $x=\alpha_2$ is an isolated zero of these functions.

THEOREM 1. *Let* (A), (B), *and* (C) *hold for some* $\alpha_1, \alpha_2, \alpha_3$ *of* $[a, b]$. *Then there exists a unique solution of* (1), (2).

The proof is similar to the corresponding theorem of [3].

We let (A′) denote the set of weaker conditions obtained from (A) by replacing $i\geqq j$ by $i>j$, and give

THEOREM 2. *Let* (A′), (B), *and* (C) *hold for some* $\alpha_1, \alpha_2, \alpha_3$ *of* $[a, b]$. *Then there exists a unique solution of* (1), (2).

Proof. We employ a transformation used in [1]. Let

$$z = F(x)y, \qquad F = \operatorname{diag}(f_i), \qquad f_i(x) = \exp\int_{\alpha_2}^{x} - a_{ii}(t)dt.$$

Then (1) becomes

$$z' = D(x)z + Fb(x), \qquad D = F'F^{-1} + FAF^{-1}. \tag{5}$$

Since $d_{ii}\equiv 0$ and, for $i\neq j$, $d_{ij}=f_i a_{ij}/f_j$, the hypotheses here imply the hypotheses of Theorem 1 for the $d_{ij}(x)$. Hence there exists a unique solution of (5) satisfying

$$z_1(\alpha_1) = \beta_1 f_1(\alpha_1),\ z_n(\alpha_3) = \beta_n f_n(\alpha_3),\ z_m(\alpha_2) = \beta_m, \qquad m = 2, \cdots, n-1.$$

This, in turn, gives a unique solution of (1), (2).

COROLLARY. *Let the* $a_{ij}(x)$ *satisfy* (A′), (B), (C) *over* $[a, \alpha_2)$, $(\alpha_2, b]$ *for some*

$\alpha_2 \in [a, b]$. *Then Theorem 2 holds without restricting* α_1, α_3 *further than requiring them to belong to* $[a, \alpha_2]$, $[\alpha_2, b]$, *respectively.*

Work on this paper was done while the author received support from NSF grant GP-1337.

References

1. W. J. Coles, Some boundary value problems for linear differential systems, Proc. Amer. Math. Soc., 14 (1963) 956–960.

2. J. B. Garner and L. P. Burton, Solutions of linear differential systems satisfying boundary conditions in the large, Proc. Amer. Math. Soc., 12 (1961) 100–106.

3. J. B. Garner, A multiple-point boundary condition for linear differential systems, this MONTHLY, 70 (1963) 850–852.

A SPECIAL POINT IN A QUADRILATERAL, OR HOW THE NINEPOINTCIRCLE BECOMES A TENPOINTCIRCLE

FR. FLOR CARTUYVELS, S.J., St. Xavier's College, Ranchi, India

Problems E 1739 and E 1740 in this MONTHLY (71(1964) 1132) draw attention to the properties of the quadrilaterals formed by the incenters or the centroids of four concyclic triangles in a concyclic quadrilateral. We extend this here to the orthocenters and the centers of the ninepoint circles. Problem E 1740 is solved in Theorem III below.

Notation. Triangle BCD has orthocenter Ω_A, center of its ninepointcircle N_A, centroid G_A; s_A is the Simson line with respect to it. Cyclic permutation of symbols gives analogical elements.

DEFINITIONS. Quadrilateral: the four vertices and the six sides joining them two by two.

Ortho-quadrilateral: the quadrilateral formed by the four orthocenters. In the same way we define a ninepoint-quadrilateral and a centro-quadrilateral.

Orthopole: see infra; co-orthopolar are elements symmetric with respect to the orthopole.

LEMMA 1a. *If four points are such that one is the orthocenter of the triangle formed by the other three, then each of the four points has this same property.*

This follows from the interplay of sides and orthogonals in the triangles. (See points A, B, C, Ω_D on the figure.)

LEMMA 1b. *The four possible triangles have the same ninepointcircle.*

This follows from the definition of the ninepointcircle.

LEMMA 1c. *These four triangles have the same circumradius.*

This circumradius is twice the radius of the unique ninepointcircle.

LEMMA 2. $\Omega_D CD\Omega_C$ *is a parallelogram. The same is true for analogical figures.*

Proof. $C\Omega_D$ and $D\Omega_C$ are obviously parallel. Moreover they are equal: trigonometry gives $C\Omega_D = AB \cot ACB$; and this is symmetric in C and D, for $ABCD$ is concyclic. Hence $C\Omega_D = D\Omega_C$.

The orthopole. In parallelograms $\Omega_D CD\Omega_C$, $\Omega_C BC\Omega_B$, $\Omega_B AB\Omega_A$, $\Omega_A DA\Omega_D$, $\Omega_C AC\Omega_A$, and $\Omega_D BD\Omega_B$ the point K is midpoint of the diagonals $D\Omega_D$, $C\Omega_C$, $B\Omega_B$, $A\Omega_A$, $A\Omega_A$, and $B\Omega_B$ respectively.

Hence K is the common midpoint of these six parallelograms. We propose to call it the *orthopole* of the quadrilateral. Pairs of points such as D and Ω_D are co-orthopolar. So are pairs of segments such as CD and $\Omega_C\Omega_D$.

The following theorem is now established:

THEOREM I. *The ortho-quadrilateral is congruent with its quadrilateral. A rotation over half a full turn around the orthopole makes the one coincide with the other.*

This theorem has the following rather long corollary. Let us call it

THEOREM II. *The following pairs of quadrilaterals are co-orthopolar with respect to the same orthopole K:*

pair (1)	$A\ B\ C\ D$	*and*	$\Omega_A\Omega_B\Omega_C\Omega_D$
pair (2)	$A\ \Omega_D C\ \Omega_B$	*and*	$\Omega_A D\ \Omega_C B$
pair (3)	$A\ \Omega_C D\ \Omega_B$	*and*	$\Omega_A C\ \Omega_C B$
pair (4)	$A\ B\ \Omega_D\Omega_C$	*and*	$\Omega_A\Omega_B D\ C$.

Moreover the two members of each pair can be made to coincide through a rotation over half a full turn, around K: hence both members have the same orthopole.

The two members of each pair are each other's ortho-quadrilateral. This follows from Lemma 1a.

This is clear enough for pair (1). For pair (2): points B, Ω_A, D, Ω_C are resp. orthocenters in triangles

$$A\Omega_D C; \quad \Omega_D C\Omega_B; \quad C\Omega_B A; \quad \Omega_B A\Omega_D.$$

The same can now be found for pairs (3) and (4).

It is relatively easy to prove that all these quadrilaterals are concyclic quadrilaterals. And all of them have the same circumradius. (Lemma 1c.)

THEOREM III. *The ortho-, ninepoint-, and centroquadrilaterals are homothetic. Hence, they are similar and concyclic.*

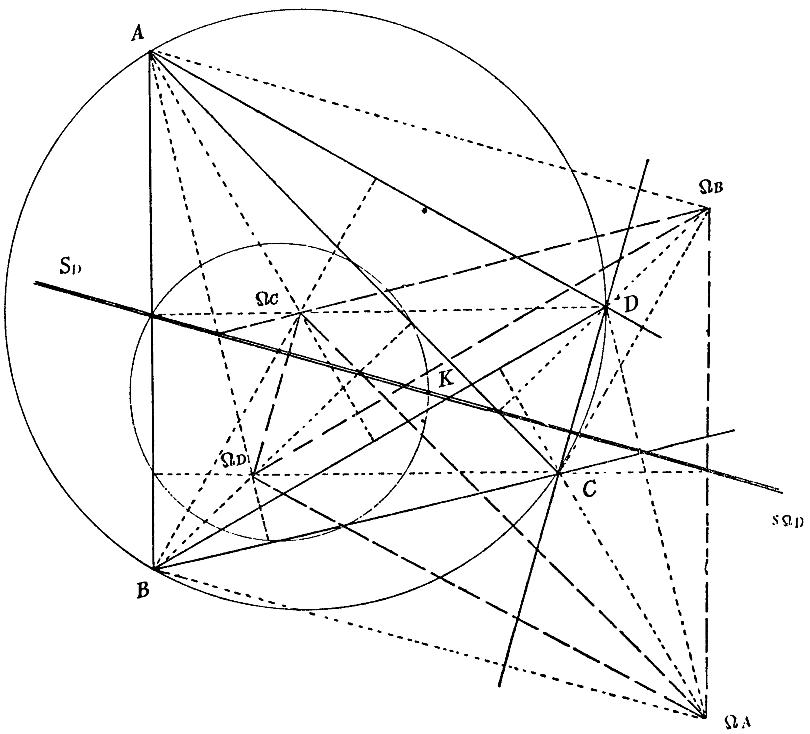
This homothety arises in the pencil of the four Euler lines. The circumcenter, which is the common point to these four Euler lines, functions as homothetic center. The ratios of similarity follow from Euler's theorem:

$$ON_A : O\Omega_A = 1/2 \quad \textit{and} \quad OG_A : O\Omega_A = 1/3.$$

Each vertex of a concyclic quadrilateral has a Simpson line with respect to the triangle formed by the three other vertices.

THEOREM IV. *These four Simpson lines are concurrent at the orthopole.*

Steiner has proved the following theorem for triangle ABC, whose orthocenter is Ω_D: the Simpson line s_D of a point D on its circumcenter passes through the midpoint of $D\Omega_D$. Hence with Lemma 2, our Theorem IV follows.



Note 1. Simpson lines are auto-co-orthopolar, i.e.: s_D is s_{Ω_D}.

Note 2. The same set of four Simpson lines are found for all the quadrilaterals of Theorem II. It can easily be seen when we complete the figure that s_D is s_B with respect to $\Omega_A D \Omega_C$, and "co-orthopolarly," s_{Ω_D} is s_{Ω_B} with respect to $A\Omega_D C$.

THEOREM V. *The four ninepointcircles have a common point, which is the orthopole.*

This follows from a known property of the ninepointcircle. The ninepointcircle can be considered as the locus of the midpoints of a radiusvector, with origin in the orthocenter, and endpoint on the circumcenter.

It follows then that K, midpoint of $\Omega_D D$ belongs to the ninepointcircle of

triangle ABC; similarly K belongs to the three other ninepointcircles. The four of them will then have K as their common point.

COROLLARY. *The orthopole is the circumcenter of the ninepoint-quadrilateral.*

The four ninepointcircles have the same radius, half the circumradius of the quadrilateral. Half the circumradius of the quadrilateral is also the circumradius to the ninepoint-quadrilateral, according to our Theorem III. Its circumcenter is therefore $\frac{1}{2}$ R (R the circumradius of the quadrilateral) away from the points N_A, N_B, N_C, N_D. And Theorem V here says that these four points are at that distance from K.

Note 1. Two by two the ninepointcircles have a second common point, the ${}_4C_2=6$ midpoints of the sides of the quadrilateral.

Note 2. K is not only common to the four ninepointcircles of the original quadrilateral $ABCD$, but also to the ninepointcircles of the ortho-quadrilateral. But the same must be said for all the pairs of quadrilaterals of Theorem II. Hence there would be 32 ninepointcircles, but Lemma 1b reduces this number to one quarter of it: for the same ninepointcircle (as on the fig.) serves for four triangles, scattered over the four pairs. Hence there will be only eight ninepointcircles, the original ones of pair (1).

To summarize we can say that the ninepointcircle becomes a tenpointcircle thanks to an eightcirclepoint.

The author wishes to thank the referee for valuable suggestions.

Reference

1. Solution of Problem E1740, Four concyclic points, this MONTHLY, 72(1965) 1026.

A NOTE ON LINEAR RECURRENT SEQUENCES MODULO m

D. W. ROBINSON, Brigham Young University

A linear recurrent sequence of integers modulo m is periodic. (See for example [1, 6].) The particular case of the Fibonacci sequence was discussed recently in this MONTHLY, (see [3, 5]). In this note we generalize the results of [3], using essentially the notation of [5].

Let (u): $u_0, u_1, \cdots, u_n, \cdots$ be the Lucas' sequence of degree r associated with the characteristic polynomial

$$f(x) = x^r - a_1x^{r-1} - \cdots - a_r,$$

where $a_1, \cdots, a_r$ are integers and $a_r \neq 0$. That is, (u) is the sequence of integers that satisfy the linear recurrence

$$u_{n+r} = a_1u_{n+r-1} + \cdots + a_ru_n$$

for $n \geqq 0$ with $u_0=0, \cdots, u_{r-2}=0, u_{r-1}=1$. (For $r=1$ it is understood that

$u_0=1$.) Also, let m be a positive integer which is relatively prime to a_r. The least positive integer k such that $u_k\equiv 0, \cdots, u_{k+r-2}\equiv 0, u_{k+r-1}\equiv 1 \pmod{m}$ is called the period $k(m)$ of (u) modulo m, and the least positive integer d such that $u_d\equiv 0, \cdots, u_{d+r-2}\equiv 0 \pmod{m}$ is called the restricted period $d(m)$ of (u) modulo m. (See for example [1, 6].)

The existence of the period, and thus the restricted period, may be established as follows. Let

$$A = \begin{pmatrix} 0 \cdots 0 & a_r \\ 1 \cdots 0 & a_{r-1} \\ \cdots & \\ 0 \cdots 1 & a_1 \end{pmatrix}$$

be the companion matrix of $f(x)$, and let U_n be the row matrix $[u_n, \cdots, u_{n+r-1}]$, $n=0, 1, \cdots$. Clearly $U_n=U_0A^n$. Since the integers modulo m form a finite system, there exist k and n such that A^{k+n} is congruent (elementwise) to A^n modulo m with $k+n>n\geqq 0$. In fact, since $\det A=(-1)^{r-1}a_r$ is a unit modulo m, A^k is congruent to the identity matrix I modulo m. Consequently, $U_k\equiv U_0 \pmod{m}$, which provides the conclusion, (see also [2, 4]).

The results of this note depend upon the following

LEMMA. *Let (u) be the Lucas' sequence of degree r associated with the polynomial $f(x)$ above, and let m be a positive integer which is relatively prime to a_r. Let $k(m)$ and $d(m)$ be the period and the restricted period of (u) modulo m, and let $s(m)$ be the order of the unit $(-1)^{r-1}a_r$ modulo m. Then $u_{d(m)+r-1}$ is a unit modulo m of order $k(m)/d(m)$, and $k(m)/d(m)$ divides $r\cdot s(m)$.*

Proof. We first show that the period $k(m)$ of (u) modulo m is precisely the order modulo m of the matrix A above. Indeed, suppose $U_k\equiv U_0 \pmod{m}$ and let M_n be the r-by-r matrix with U_{n+i-1} for its ith row. Since $U_{k+i}\equiv U_i \pmod{m}$, it follows that $M_k=M_0A^k\equiv M_0 \pmod{m}$. But, since $\det M_0=(-1)^{r-1}$, M_0 is a unit and $A^k\equiv I \pmod{m}$. That is, $A^k\equiv I \pmod{m}$ if and only if $U_k\equiv U_0 \pmod{m}$. Consequently, $k(m)$ is the order of A modulo m.

We may prove in a similar manner that $A^d\equiv tI \pmod{m}$ if and only if $U_d\equiv tU_0 \pmod{m}$, where t is some integer. Thus the restricted period $d(m)$ generates the ideal of all d such that A^d is a scalar matrix modulo m. In particular, $d(m)\mid k(m)$. Also, since $U_0=[0, \cdots, 0, 1]$, if $U_d\equiv tU_0 \pmod{m}$ then $t\equiv u_{d+r-1} \pmod{m}$. Therefore,

$$A^{d(m)} \equiv u_{d(m)+r-1}I \pmod{m},$$

and it follows that $u_{d(m)+r-1}$ is a unit modulo m of order $k(m)/d(m)$. Finally, if $s(m)$ is the order of $\det A=(-1)^{r-1}a_r$ modulo m, then

$$1 \equiv (\det A)^{s(m)\cdot d(m)} \equiv (\det A^{d(m)})^{s(m)} \equiv (u_{d(m)+r-1})^{r\cdot s(m)} \pmod{m}.$$

That is, the order of $u_{d(m)+r-1}$ modulo m divides $r\cdot s(m)$.

We now extend the first theorem of [**3**].

THEOREM. *Let the notation and conditions be as in the lemma. If p is a prime such that $p\nmid a_r$, then $k(p^e)=k(p)$ implies $d(p^e)=d(p)$.*

Proof. Since $A^{k(p)}=I+pB$ for some matrix B, it is clear that

$$A^{p^{e-1}k(p)}\equiv I \pmod{p^e}.$$

That is, $k(p^e)\mid p^{e-1}k(p)$. But it is obvious that $k(p)\mid k(p^e)$. Hence, $k(p^e)/k(p)$ is some nonnegative power of p. Similarly, $d(p^e)/d(p)$ is a nonnegative power of p. Also, since the unit group of integers modulo p is of order $p-1$, it follows by the first conclusion of the lemma that $k(p)/d(p)$ divides $p-1$. Thus, $p\nmid k(p)/d(p)$ and the theorem is a consequence of the fact that

$$\frac{d(p^e)}{d(p)}\cdot\frac{k(p^e)}{d(p^e)}=\frac{k(p^e)}{k(p)}\cdot\frac{k(p)}{d(p)}.$$

A corollary of these results is that if $p\nmid a_r$ and $k(p^e)=k(p)$, then $u_{d(p)+r-1}$ has the same order modulo p and modulo p^e. Indeed, we have the following generalization of the second theorem of [**3**].

COROLLARY 1. *Let the notation and conditions be as in the lemma. If p is a prime such that $p\nmid a_r$ and e is a positive integer such that $d(p^e)=d(p)$, then $k(p^e)=k(p)$ if and only if $u_{d(p)+r-1}$ has the same order modulo p and modulo p^e.*

Furthermore, by the final conclusion of the lemma, it is immediate that we also have the following

COROLLARY 2. *Let the notation and conditions be as in the lemma and let $a_r=\pm 1$. If p is a prime such that $p\nmid 2r$, then $k(p^e)/d(p^e)=k(p)/d(p)$ for every positive integer e.*

If $a_r=(-1)^{r-1}$ in this corollary, then it is sufficient to require only $p\nmid r$. This condition is necessary, however, as shown by the following example. If $f(x)=x^3-x^2-x-1$ and $p=3$, then $k(3)=d(3)=d(3^2)=13$, but $k(3^2)=39$.

This paper was written while the author was on sabbatical leave at the California Institute of Technology, where he was supported in part by the National Science Foundation.

References

1. R. D. Carmichael, On sequences of integers defined by recurrence relations, Quart. J. Math., 48 (1920) 343–372.

2. E. C. Dade, D. W. Robinson, O. Taussky, and M. Ward, Divisors of recurrent sequences, J. Reine Angew. Math., 214/215 (1964) 180–183.

3. S. E. Mamangakis, Remarks on the Fibonacci series modulo m, this MONTHLY, 68 (1961) 648–649.

4. D. W. Robinson, The Fibonacci matrix modulo m, The Fibonacci Quart., 1 (1963) 29–36.

5. D. D. Wall, Fibonacci series modulo m, this MONTHLY, 67 (1960) 525–532.

6. M. Ward, The characteristic number of a sequence satisfying a linear recursion relation, Trans. Amer. Math. Soc., 33 (1931) 153–165.

$$d_{\delta\gamma}e^{\delta}du^{\gamma} = 0. \tag{4.1}$$

According to the definition given in [1, 957], equation (4.1) represents the asymptotic line of the vector field e^{α}. Consequently, we have

THEOREM 4.1. *The derived vector of the unit normal along the asymptotic line of a vector field in a surface is orthogonal to the curve of the vector field.*

The above is a generalization of the classical theorem that the derived vector of the unit normal along a curve of a surface is orthogonal to the curve provided it is an asymptotic line of the surface.

References

1. T. K. Pan, Normal curvature of a vector field, Amer. J. Math., 74 (1952) 955–966.
2. C. E. Springer, Rectilinear congruences whose developables intersect a surface in its lines of curvature, Bull. Amer. Math. Soc., 51 (1945) 990–996.

SYMMETRY IN GEOMETRIC COMPLEXES

P. H. DOYLE, Michigan State University

A complex K will always mean a finite geometric simplicial complex. For a complex K we shall let $G(K)$ be the group of homeomorphism of K onto K. For a vertex v of K, $\mathrm{St}(v)$ is the star of v.

Following [2] we define for a simplicial complex K the set $I(K)$ consisting of all points x in K such that for each open set U in K containing x there exists a homeomorphism k in $G(K)$ such that $k(K-U)\subset U$; $I(K)$ is called the invert set of K.

LEMMA 0. *Let K be a complex. If x and y are points lying in the interior of a k-simplex σ^k of K, then there exists a homeomorphism k in $G(K)$ and $k(x)=y$. Further, k may be chosen so that $k|K-\mathrm{St}\,\sigma^k$ is the identity homeomorphism and each simplex is invariant in K.*

Proof. If $\dim K=0$, the lemma is clear. Suppose we have established Lemma 0 for all simplicial complexes K for which $\dim K<n$. We then assume dimension $K=n$ and let $K^{(n-1)}$ be the $(n-1)$-skeleton of K.

If x and y are points interior to an n-simplex of K the homeomorphism of Lemma 0 is at hand. Otherwise we may apply the inductive hypothesis to $K^{(n-1)}$ and extend a given homeomorphism on $K^{(n-1)}$ to all K. This completes the proof.

A topological space X is rigid if the only homeomorphism of X onto X is the identity homeomorphism.

COROLLARY 1. *The only rigid complex is a point.*

COROLLARY 2. *If K is a finite complex and the conditions of Lemma 0 are met, x and y are equivalently embedded in K.*

THEOREM 1. *Let K be a complex with a fixed triangulation T. Then $I(K)$ is a subcomplex of K with respect to T.*

Proof. If $I(K)=\square$, there is nothing to prove. Then we assume $I(K)\neq\square$. Suppose that x is a point in $I(K)$ that lies interior to a simplex σ^k of K under T, By Lemma 0 and [2] it follows that $\operatorname{Int}\sigma^k \subset I(K)$. But since $I(K)$ is closed. $\sigma^k \subset I(K)$. It follows that $I(K)$ is a union of closed simplices of K under T and Theorem 1 is proved.

In what follows we need the notion of a suspension ring. If K is a finite simplicial complex in E^n, euclidean n-space, imagine $E^n \subset E^{n+1}$. If a_1 and b_1 are points, one in each component of $E^{n+1}-E^n$, then the suspension of K, $S(K)$, consists of the union of all segments a_1x and b_1x for x in K. If now we take the suspension of $S(K)$ using points a_2, b_2 in E^{n+2}, the resulting set $D(K)$ is the double suspension of K with $a_1a_2 \cup b_2a_1 \cup a_2b_1 \cup b_2b_1$ as suspension ring.

THEOREM 2. *If K is a complex then $I(K)$ is null, a point, or a finite simplicial sphere that is a subcomplex of K under arbitrary simplicial subdivision.*

Proof. If $I(K)\neq\square$, then by [2] we observe that $I(K)$ must be a sphere. By Theorem 1 $I(K)$ is always a subcomplex of K.

COROLLARY. *Let K be a finite polyhedron. Then there is a simplicial subdivision of K containing $I(K)$ as a subcomplex; namely, each simplicial subdivision of K.*

THEOREM 3. *If K is an n-complex which is not a point, let $I(K)\neq\square$. If $K/I(K)$ is invertible, then K is a sphere or a cell if K is a manifold or has a free $(n-1)$-face.*

Proof. If $K/I(K)$ is a point, K is a sphere. So let $K/I(K)$ be different from a point.

This implies that $I(K)\neq K$ and so if T is a triangulation of K, $I(K)$ fails to contain a principal simplex σ^k of K under T. But then $K/I(K)$ must be an n-sphere. Note that $K-I(K)=E^n$. It follows that $I(K)$ must contain an $(n-1)$-simplex of K. Thus $I(K)$ is an $(n-1)$-sphere.

Let σ^{n-1} be an $(n-1)$-simplex in $I(K)$ and p a point of $\operatorname{Int}\sigma^{n-1}$. Since σ^{n-1} must lie on a face of some n-simplex σ^n in K, it follows that $I(K)$ is locally collared in K. (σ^{n-1} can't lie as a face of 2 n-simplices, for by the proof of Lemma 0, $K=S^n$). Then K must be an n-manifold with boundary and by [2], K is an n-cell.

THEOREM 4. *Let K be a simplicial complex and $D(K)$ its double suspension. If R is the suspension ring, then $R\subset I(D(K))$.*

Proof. The vertices of the second suspension of K lie in $I(D(K))$ and R is homogeneously embedded in $D(K)$.

One can ask of $D(K)$ in Theorem 4, if $R\neq I$ $(D(K)$ what is $I(D(K))$? In this case it is clear that $D(K)-R$ contains a point p of $I(D(K))$. Whence if $S(K)$ is the first suspension of K we may assume that p lies in $S(K)-R$. But then we must have a point of K in $I(D(K))$. In this case $I(D(K))$ is at least a 2-sphere.

THEOREM 5. *Let K be a triangulated compact n-manifold and $D(K)$ its double suspension. Then $I(D(K))=R$, unless $D(K)$ is a sphere. Further, if $I(D(K))=R$ then $D(K)/I(D(K))$ is locally an $(n+2)$-manifold except at one point.*

Proof. If $I(D(K))\neq R$, $D(K)$ is locally euclidean and so a sphere.

If $D(K)/I(D(K))$ were locally euclidean, then by Theorem 3, $D(K)$ would be a sphere or a cell. But we assume that $D(K)$ is neither.

Construction. Theorem 5 provides a scheme for constructing complexes with precisely one invert point. If K is a nonsimply connected compact n-manifold, then $D(K)/R$ is precisely this type of complex.

THEOREM 6. *Let K be a connected complex of* dim n *and* dim $I(K)=n-1>0$. *Then K is an n-cell.*

Proof. $I(K)=S^{n-1}$. Since K contains an n-simplex, we note that each $(n-1)$-simplex in $I(K)$ must be on the face of an n-simplex and precisely one. Thus $I(K)$ is locally collared. Whence $K-I(K)$ is locally euclidean. It follows that K is an n-cell.

A point x in K is a point of continuous invertibility if x lies in $I(K)$ and inverting homeomorphisms for neighborhoods of x are isotopic to the identity.

THEOREM 7. *All points of a suspension ring are points of continuous invertibility.*

Proof. If R is the ring of $D(K)$, let v_1, v_2 be vertices of the 2nd suspension. If U is an open set containing v_1 and if w_1, w_2 are vertices of the first suspension, there exists an isotopy k_t pushing $D(K)-U$ into a preassigned open set containing v_2 with closure not meeting w_1, w_2. There is then an isotopy pushing $k_1(D(K)-U)$ into any preassigned neighborhood of w_1. Then a third isotopy takes $g_1k_1(D(K)-U)$ into U.

If K is a finite complex, $H(K)$ is the number of equivalence classes of points under the relation of equivalent under elements of $G(K)$.

THEOREM 8. $H(S(K))\leqq H(K)+1$ *and* $H(D(K))\leqq H(S(K))$, *where $S(K)$ is the suspension of K.*

In [1] a polyhedral homotopy 5-sphere is constructed by a double suspension of a 3-dimensional Poincaré manifold. The question of its invertibility is as yet unsettled.

The general problems of classifying geometric complexes with nonvoid invert sets present some interesting challenges. Properties of such complexes are being investigated by Klassen [3]. In the case of a 1-dimensional complex K with exactly one invert point, it is a simple matter to prove the following.

THEOREM 9. *A* 1-*complex K with exactly* 1 *point in $I(K)$ is a set of n simple closed curves meeting in a single point, but otherwise disjoint in pairs (an n-leafed rose).*

References

1. M. L. Curtis and E. C. Zeeman, On the polyhedral Schoenflies theorem, Proc. Amer. Math. Soc., 2 (1960) 269–279.
2. P. H. Doyle, J. G. Hocking and R. P. Osborne, Local Invertibility, Fund. Math., 53 (1963) 887–888.
3. V. Klassen, Thesis, Virginia Polytech. Inst., (1964).

NECESSARY AND SUFFICIENT CONDITION FOR A CONVEX SET TO BE CLOSED

HUBERT HALKIN, Bell Telephone Laboratories, Whippany, N. J.

Introduction. The existence of solutions of many optimization problems is closely related to the closure of some convex sets in an n-dimensional Euclidean space E^n. (LaSalle [1], Neustadt [2], Halkin [3].) Suppose that A is a bounded, non-empty convex set in E^n. For any nonzero vector p in E^n we define $H(p)$ to be the supporting hyperplane of A which is normal to the vector p and we let $G(p)=H(p)\cap A$. In this note we shall derive a necessary and sufficient condition for the set A to be closed in terms of some properties of the sets $G(p)$. We see

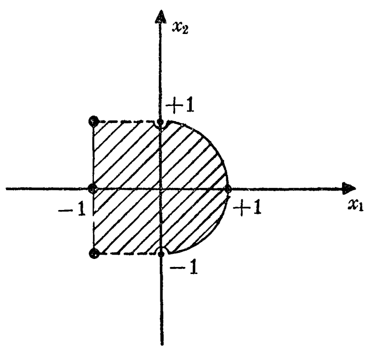


FIG. 1.—Convex set A.

immediately that a necessary condition for the set A to be closed is that the sets $G(p)$ are closed and non-empty for any nonzero vector p. The preceding condition is not sufficient, however, as we can see from the example in E^2 given by Figure 1. In that example we have $A=A_1\cup A_2\cup A_3$ where

$$A_1 = \{(x_1, x_2): (x_1)^2 + (x_2)^2 \leqq 1, \ |x_2| < 1\}$$

$$A_2 = \{(x_1, x_2): |x_2| < 1, x_1 \in [-1, 0]\}$$

$$A_3 = \{(x_1, x_2): x_1 = -1, x_2 \in [-1, +1]\}.$$

The preceding example indicates clearly that we must add some requirements on the continuity of the dependence of the set $G(p)$ upon the vector p in order to have sufficient conditions for the set A to be closed. We define the distance $d(G(p'), G(p''))$ between two sets $G(p')$ and $G(p'')$ as the infimum of $|a'-a''|$ where $a'\in G(p')$ and $a''\in G(p'')$. In this note we shall prove the following result:

THEOREM I. *A necessary and sufficient condition for the bounded, nonempty, convex set A to be closed is that* (i) *for every nonzero vector p the set $G(p)$ is closed and nonempty*; (ii) *if $p_1, p_2, \cdots$ is a sequence of nonzero vectors converging to some nonzero vector p then* $\lim_{i=\infty} d(G(p_i), G(p))=0$.

Proof. We shall first prove that the condition is necessary. Part (i) is obtained immediately. Let $p_1, p_2, \cdots$ be a sequence of nonzero vectors converging to the nonzero vector p. For every $i=1, 2, \cdots$ let a_i be a point in $G(p_i)$ nearest to the set $G(p)$; this point a_i exists since the sets $G(p_i)$ and $G(p)$ are closed and bounded. Let d_i be the distance from a_i to the set $G(p)$. We conclude by proving, by contradiction, that $\lim_{i=\infty} d_i=0$. Let us assume that there is an $\eta>0$ and subsequence $k_1, k_2, \cdots$ of $1, 2, \cdots$ such that $d_{k_i}\geqq\eta$ for all $i=1, 2, \cdots$. The sequence of points $a_{k_1}, a_{k_2}, \cdots$ is bounded, hence, it has a subsequence $a_{m_1}, a_{m_2}, \cdots$ converging to some point a which does not belong to $G(p)$. We have $a\in A$ since the set A is closed. We shall use a dot to indicate the scalar product of two vectors. For every $x\in A$ we have

$$a\cdot p = \lim_{i=\infty} (a_{m_i}\cdot p_{m_i}) \geqq \lim_{i=\infty} (x\cdot p_{m_i}) = x\cdot p$$

which implies that $a\in H(p)$ and contradicts the assumption $a\notin G(p)$.

Let us now prove that the condition is sufficient. Let $a\in\partial A$. We shall prove that $a\in A$. We know that $a\in H(p)$ for some nonzero vector p. Let us prove, by contradiction, that $a\in G(p)$. If $a\notin G(p)$ let b be the point of $G(p)$ closest to a; the point b exists since the set $G(p)$ is closed. Moreover $b\neq a$. For every $\epsilon>0$ let $p_\epsilon=p+\epsilon(a-b)$. There exists an $\bar{\epsilon}>0$ such that

$$d(G(p), G(p_{\bar{\epsilon}})) \leqq \frac{|a-b|}{2}.$$

Let $a^*\in G(p_{\bar{\epsilon}})$ and $b^*\in G(p)$ such that $|a^*-b^*|\leqq|a-b|/2$. We have $(p+\bar{\epsilon}(a-b))\cdot a\leqq(p+\bar{\epsilon}(a-b))\cdot a^*$ and $p\cdot a^*\leqq p\cdot a$; hence $(a-b)\cdot(a^*-a)\geqq 0$. Since the vector $a-b$ is an outward normal to a supporting hyperplane of the convex set $G(p)$ at the point b and since $b^*\in G(p)$ we have also $(a-b)\cdot(b-b^*)\geqq 0$. From the last two inequalities we obtain immediately $(a-b)\cdot(a^*-b^*)\geqq(a-b)\cdot(a-b)$ which implies $|a^*-b^*|\geqq|a-b|$ and contradicts the relation $|a^*-b^*|\leqq|a-b|/2$ assumed earlier.

The author is presently at the University of California, San Diego. Research sponsored by the Air Force Office of Scientific Research, US Air Force, under AFOSR Grant No. 1039-66.

References

1. J. P. LaSalle, The time optimal control problem, in "Contributions to the theory of nonlinear oscillations," vol. V, Princeton University Press, 1960, pp. 1–24.
2. L. W. Neustadt, The existence of optimal control in the absence of convexity conditions, J. Math. Anal. Appl., 7 (1963) 110–117.
3. H. Halkin, On a generalization of a theorem of Lyapunov, J. Math. Anal. Appl., 10 (1965) 325–329.

A NOTE ON MATRICES WITH ZERO TRACE

FERGUS GAINES, California Institute of Technology

1. It is known that an $n\times n$ matrix A with elements in a field F can be written as a commutator $XY-YX$ over F, if and only if the trace of A is zero. This result was proved by Shoda [1] in the case that the field has characteristic zero and was extended by Albert and Muckenhoupt [2] to fields of arbitrary characteristic. In this note we consider Shoda's result when F is the real or complex numbers and we also derive some results when A is hermitian or skew-hermitian.

The author wishes to express his thanks to Dr. Olga Taussky for suggesting to him the results contained in the corollaries to Theorem 1.

2. We shall make use of the following two results of W. V. Parker [3].

LEMMA 1. *If A is an $n\times n$ matrix of complex numbers and trace A is zero, then there exists a unitary matrix U so that UAU^* has all its main diagonal elements equal to zero.*

LEMMA 2. *If A is a real $n\times n$ matrix with trace zero, then there exists a real orthogonal matrix T so that TAT^t has zero main diagonal.*

Our main result is the following:

THEOREM 1. *If A is an $n\times n$ complex matrix and trace A is zero, then there exist matrices X and Y so that $A=XY-YX$, where X is hermitian and Y has trace zero.*

Proof. By Lemma 1 we can find a unitary matrix U so that $UAU^*=B=(b_{ij})$ has zero diagonal. Let $D=\operatorname{diag}(d_1, d_2, \cdots, d_n)$ where the d_i are real and distinct, and let $Y_1=(y_{ij})$ where $y_{ij}=b_{ij}/(d_i-d_j)$ when $i\neq j$ and $y_{ii}=0$, $i, j=1, 2, \cdots, n$. Then $B=DY_1-Y_1D$ and thus $A=XY-YX$ where $X=U^*DU$, $Y=U^*Y_1U$, and we see that $X^*=X$ and trace $Y=0$.

COROLLARY 1. *If, in addition, A is hermitian, then it can be written as $XY-YX$, where X is hermitian and Y is skew-hermitian.*

COROLLARY 2. *If A is skew-hermitian with trace zero, it can be written as $XY-YX$ where both X and Y are hermitian.*

REMARK. If A is a real matrix with trace zero then Theorem 1 and its corollaries hold if we replace "hermitian" and "skew-hermitian" by "symmetric" and "skew-symmetric," respectively. Lemma 2 is used to prove these facts.

If, in Corollary 1, we replace X by $B = (1/\sqrt{2})(X - Y)$ and Y by $B^* = (1/\sqrt{2})(X+Y)$, we get the following theorem due to R. C. Thompson [4].

THEOREM 2. *If A is a hermitian matrix and trace $A=0$, then A can be written as BB^*-B^*B.*

It is also true that Corollary 1 follows from Thompson's result on replacing B by $(1/\sqrt{2})(X-Y)$ where $X=(1/\sqrt{2})(B^*+B)$ and $Y=(1/\sqrt{2})(B^*-B)$.

This work was carried out (in part) under a grant from the National Science Foundation.

References

1. K. Shoda, Einige Sätze über Matrizen, Japan J. Math., 13 (1936) 361–365.
2. A. A. Albert and B. Muckenhoupt, On matrices of trace zero, Michigan Math. J., 4 (1957) 1–3.
3. W. V. Parker, Sets of complex numbers associated with a matrix, Duke Math. J., 15 (1948) 711–715.
4. R. C. Thompson, On matrix commutators, J. Washington Acad. Sci., 48 (1958) 306–307.

SOLUTIONS OF $x^4+y^4=z^4$ IN 2×2 INTEGRAL MATRICES

R. Z. DOMIATY, Technische Hochschule, Graz

A. Aigner [1] investigated solutions of $x^4+y^4=z^4$ in quadratic domains. In this note, solutions are found in the ring Γ of 2×2 matrices with integer elements. Let 0 and I be the zero and identity matrices respectively.

THEOREM. *There exist solutions of $A^4+B^4=C^4$, where A, B and C are in Γ and $A^4\neq 0$, $B^4\neq 0$, $C^4\neq 0$.*

Proof. Set

$$A = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & c \\ d & 0 \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & e \\ f & 0 \end{bmatrix}.$$

Then $A^4=(ab)^2I$, and there are similar expressions for B^4 and C^4. Thus, $A^4+B^4=C^4$ if and only if $(ab)^2+(cd)^2=(ef)^2$. But using the well-known solution to the Diophantine equation $x^2+y^2=z^2$, we can set $b=d=f=1$, and $a=2mn$, $c=m^2-n^2$, $e=m^2+n^2$, obtaining

$$\begin{bmatrix} 0 & 2mn \\ 1 & 0 \end{bmatrix}^4 + \begin{bmatrix} 0 & m^2-n^2 \\ 1 & 0 \end{bmatrix}^4 = \begin{bmatrix} 0 & m^2+n^2 \\ 1 & 0 \end{bmatrix}^4.$$

Reference

1. A. Aigner, Über die Möglichkeit von $x^4+y^4=z^4$ in quadratischen Körpern, Jber. Deutsch. Math. Verein., 43 (1934) 226–229.

A NOTE ON HIGHER COMMUTATORS OF BOUNDED NILPOTENCE

T. P. KEZLAN, University of Texas

A well-known theorem of Kaplansky, Herstein, and Kleinfeld states that if R is a ring for which there exists a fixed positive integer q such that $(xy-yx)^q=0$ for all x, y in R, then the nilpotent elements of R form an ideal [4, p. 29; 5]. It follows quite easily in fact that the commutator ideal of R is nil. For x, y in a ring R we denote the commutator $[x, y]=xy-yx$ by $e_1(x, y)$; the continued kth commutator is defined by $e_k(x, y)=[e_{k-1}(x, y), y]$ for $k>1$. It is the purpose of this note to generalize the theorem stated above as follows:

THEOREM. *If R is a ring for which there exist fixed positive integers k, n, and q such that $(e_k(x, y^n))^q=0$ for all x, y in R, then the commutator ideal of R is nil.*

We first give two lemmas, the first of which is Theorem 4.4 of [2] and whose proof is omitted here.

LEMMA 1. *Let R be a semi-simple ring such that for all x, y in R there exist integers $k=k(x, y)$ and $n=n(x, y)$ such that $e_k(x, y^n)=0$. Then R is commutative.*

LEMMA 2. *Let R be a semi-simple ring such that for all x, y in R there exist integers $k=k(x, y)$, $n=n(x, y)$, and $q=q(x, y)$ such that $(e_k(x, y^n))^q=0$. Then R is commutative.*

Proof. First if R is a division ring, then since R has no nonzero nilpotent elements, R satisfies the hypothesis of Lemma 1 and hence is commutative.

Now suppose that R is primitive. If R is not a division ring, then the ring Δ_2 of 2×2-matrices over some division ring Δ is a homomorphic image of a subring of R and thus inherits the property in the hypothesis of the lemma. But this cannot happen, as is easily seen by taking

$$x=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\qquad y=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}.$$

Hence R must be a division ring and therefore commutative by the first part of the proof.

Finally if R is semi-simple, then R is a subdirect sum of primitive rings R_α. Each R_α, as a homomorphic image of R, satisfies the hypothesis of the lemma and hence is commutative by the above discussion. Thus R, being a subdirect sum of commutative rings, is commutative.

Proof of the theorem. By a result of Herstein [3] it suffices to prove the theorem for the case in which R is an algebra over a field, so assume that this is the case.

Let $c\in C(R)$, the commutator ideal of R. Then for some positive integer m and each i such that $1\leqq i\leqq m$, there exist elements r_i, s_i, u_i, v_i, x_i, and y_i in R and integers q_i such that, letting $z_i=x_iy_i-y_ix_i$,

$$c = \sum_{i=1}^{m} (q_i z_i + r_i z_i + z_i s_i + u_i z_i v_i).$$

Let S be the subalgebra of R generated by all r_i, s_i, u_i, v_i, x_i, and y_i. S is finitely generated and satisfies the polynomial identity $(e_k(x, y^n))^q = 0$; thus by a result of Amitsur [1], the Jacobson radical $J(S)$ is nil. The semi-simple ring $S/J(S)$ satisfies the identity $(e_k(x, y^n))^q = 0$; hence by Lemma 2, $S/J(S)$ is commutative. Therefore all commutators of S are in $J(S)$ whence $C(S) \subseteq J(S)$. But now we have $c \in C(S) \subseteq J(S)$ with $J(S)$ nil. Thus c is nilpotent, and since $c \in C(R)$ was arbitrary, $C(R)$ is nil.

References

1. S. A. Amitsur, A generalization of Hilbert's Nullstellensatz, Proc. Amer. Math. Soc., 8 (1957) 649–656.
2. M. P. Drazin, Engel rings and a result of Herstein and Kaplansky, Amer. J. Math., 77 (1955) 895–913.
3. I. N. Herstein, A remark on rings and algebras, Mich. Math. J., 10 (1963) 269–272.
4. ———, Theory of Rings, University of Chicago Lecture Notes, 1961.
5. E. Kleinfeld, Simple alternative rings, Ann. Math., 58 (1953) 544–547.

A NOTE ON ORTHOSTOCHASTIC MATRICES

D. Ž. Djoković, Belgrade, Yugoslavia

Let $\tilde{M}_n(M_n)$ be the set of all n by n matrices with complex (real) elements. By $\tilde{U}_n(U_n)$ we shall denote the set of all n by n unitary matrices with complex (real) elements. A matrix $A = (a_{ij}) \in M_n$ is said to be doubly stochastic (d.s.) if and only if

$$a_{ij} \geqq 0, \qquad \sum_{j=1}^{n} a_{ij} = 1, \qquad \sum_{i=1}^{n} a_{ij} = 1, \qquad 1 \leqq i, j \leqq n.$$

Let S_n be the set of all d.s. n by n matrices. We define the mapping f: $\tilde{U}_n \to S_n$ by

$$U = (u_{ij}) \in \tilde{U}_n \Rightarrow f(U) = (|u_{ij}|^2).$$

It is known [1] that $f(U_n)$ is properly contained in S_n for $n > 2$. The same is true for $f(\tilde{U}_n)$. Let us put $\tilde{O}_n = f(\tilde{U}_n)$ and $O_n = f(U_n)$. If $O \in \tilde{O}_n$ we say that O is c-orthostochastic. If $O \in O_n$ we say that O is orthostochastic (or r-orthostochastic).

The set $\tilde{M}_n$ is a normed linear space with respect to the norm

$$\|M\| = \operatorname{tr}(MM^*) \qquad (M \in \tilde{M}_n).$$

If $M \in M_n \subset \tilde{M}_n$ this norm coincides with the euclidean norm.

In [2, p. 332] Mirsky proposed the following problem:

Is O_n everywhere dense, with respect to the euclidean norm, in S_n?

It is not difficult to see that the answer is negative. Moreover we shall prove that the answer is negative also if O_n is replaced by $\tilde{O}_n$. First we prove that $\tilde{U}_n$ is closed in $\tilde{M}_n$. Indeed, let $\lim_{k\to\infty} A_k = A$, $A_k \in \tilde{U}_n$ $(k=1, 2, \cdots)$. Then we have also $\lim_{k\to\infty} A_k^* = A^*$ and

$$AA^* = (\lim_{k\to\infty} A_k)(\lim_{k\to\infty} A_k^*) = \lim_{k\to\infty} (A_k A_k^*) = I,$$

i.e. $A \in \tilde{U}_n$. Further, $\tilde{O}_n$ is compact, for $\tilde{U}_n$ is compact and f is continuous. Hence, $\tilde{O}_n$ is not everywhere dense in S_n.

It is also possible to prove that $\tilde{O}_n$ and O_n are arcwise connected.

I wish to express my thanks to Prof. L. Mirsky for encouragement in preparing this note.

References

1. M. Marcus, Some properties and applications of doubly stochastic matrices, this MONTHLY, 67 (1960) 215–221.

2. L. Mirsky, Results and problems in the theory of doubly stochastic matrices, Z. Wahrscheinlichkeitstheorie, 1 (1963) 319–334.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

Address all correspondence to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

PATHS OF MINIMAL LENGTH WITHIN A CUBE

R. A. JACOBSON, Houghton College, AND K. L. YOCOM, South Dakota State University

It is often customary to discuss extremum problems by employing calculus in an entirely analytic context. However, there are certain classes of problems, not often mentioned, that are more readily attacked from a different viewpoint [1–5]. In this note, we consider several problems of this type, develop a rather general method of attack, and point out some of the difficulties which might still be encountered. In particular, we shall seek to find closed paths of minimal length within a cube that touch certain faces in specified or arbitrary sequences.

Let Q be the cube, Figure 1a, with faces A, B, C, D, E, F, lying in planes $x=0$, $x=2$, $y=0$, $y=2$, $z=0$, $z=2$, respectively; and consider the following:

Problem 1. Find the shortest closed path which begins at point $p(0, 1, 1)$ on face A and touches the faces in the sequences (a) C-D-B-F-E; (b) C-B-F-D-E; (c) C-D-F-E-B.

Problem 2. Find the shortest closed path which begins at point p and touches each face of the cube at least once.

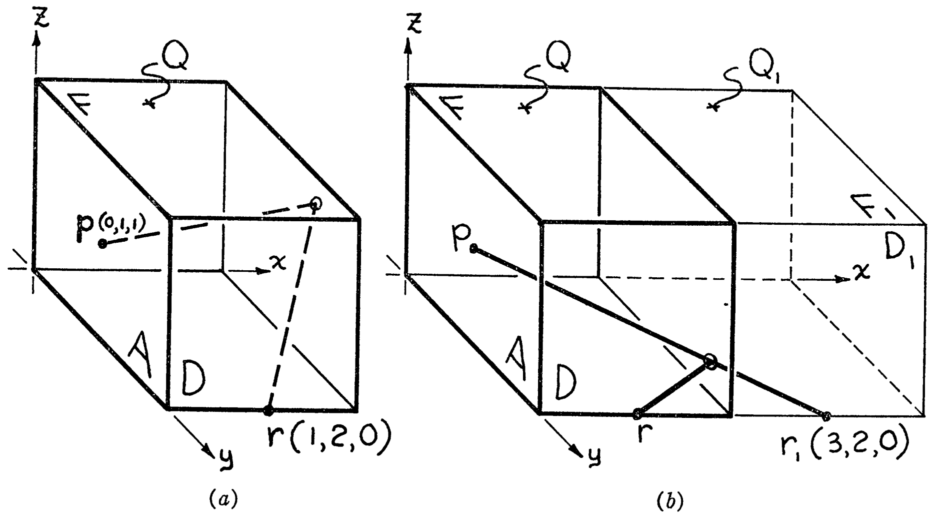


FIG. 1

In order to emphasize the complexity of the problem when attacked with calculus, we pause to consider problem 1a in an analytic setting. Letting $(x_1, 0, z_1)$, $(x_2, 2, z_2)$, $(2, y_3, z_3)$, $(x_4, y_4, 2)$, $(x_5, y_5, 0)$ denote five points on faces C, D, B, F, E, respectively, one would seek to minimize the function

$$\begin{aligned} L = &\sqrt{(x_1-0)^2+(0-1)^2+(z_1-1)^2}+\sqrt{(x_2-x_1)^2+(2-0)^2+(z_2-z_1)^2} \\ &+\sqrt{(2-x_2)^2+(y_3-2)^2+(z_3-z_2)^2}+\sqrt{(x_4-2)^2+(y_4-y_3)^2+(2-z_3)^2} \\ &+\sqrt{(x_5-x_4)^2+(y_5-y_4)^2+(0-2)^2}+\sqrt{(0-x_5)^2+(1-y_5)^2+(1-0)^2} \end{aligned}$$

in ten variables where the domain for each variable is $[0, 2]$. Proceeding formally, one would seek values for the ten variables which would simultaneously render each of the ten first partial derivatives equal to zero. Such an attack, as well as being formidable, is indeed hopeless since no such set of values exist for the specified domain. This fact becomes apparent later in the note.

Before discussing the proposed problems, we find it beneficial to consider a somewhat simpler problem. In particular let us connect the points $p(0, 1, 1)$ and $r(1, 2, 0)$ with the path of minimum length that touches face B, figure 1a.

Solution. Let q be a point on face B. Reflecting cube Q in face B we obtain the cube Q_1, figure 1b, with faces $A_1, B_1, C_1, D_1, E_1, F_1$ lying in planes $x=4$, $x=2$, $y=0$, $y=2$, $z=0$, $z=2$, respectively. Since length is invariant under reflection,

the length of the paths p-q-r and p-q-r_1 (3, 2, 0) are identical. Thus the straight line, $\overline{pr_1}$, reflected back into cube Q will indicate the shortest route. In particular, the path connects the points p, $q(2, 5/3, 1/3)$, and r.

The original problems will be handled in a similar manner; that is, cube Q will be successively reflected through appropriate faces in a given sequence, thus generating a particular polyhedron. We then seek the path of minimum length lying entirely within the polyhedron that joins the original point and its final image.

Problem 1a, *Solution.* Reflecting Q consecutively through faces C-D-B-F-E leads to the polyhedron composed of cubes Q, Q_1, Q_2, Q_3, Q_4, Q_5, figure 2a; where the final image of p is point $p_5(4, -3, 5)$ on face A_5.

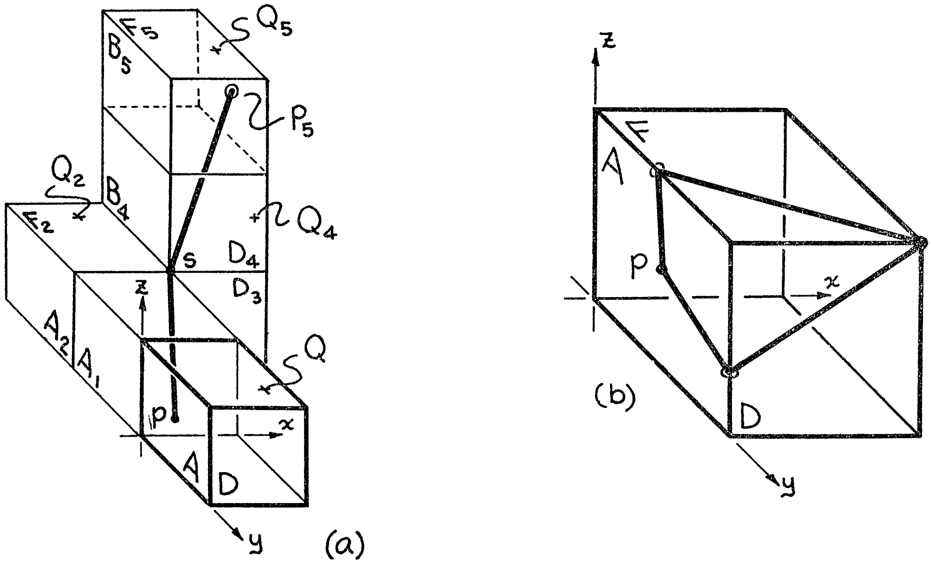


FIG. 2

It is apparent that the shortest path connecting p and p_5 within the polyhedron is composed of the two line segments, each of length $\sqrt{14}$ units, from p to s (2, −2, 2) and from s to p_5. This path when reflected back into the original cube connects the points: p; (2/3, 0, 4/3) on C; (2, 2, 2) on D, B, and F; (2/3 4/3, 0) on E; and p; in that sequence, figure 2b.

Problem 1b, *Solution.* Reflecting Q consecutively through the faces C-B-F-D-E, figure 3a; we seek the shortest path within the region connecting p and $p_5(4, -3, 5)$.

Although the final solution is not immediately obvious, one can see that the path will consist of two line segments lying in plane H_1 and H_2 determined by points p, s, $t(2, -4, 2)$ and p_5, s, t, respectively. Revolving plane H_1 about

line $\overline{st}$ until it is parallel to H_2, we find that the straight line joining p and p_5, figure 3b, gives the shortest distance, $\sqrt{34+2\sqrt{65}}$ units. When reflected back into cube Q this line becomes the path connecting points: p; $[(5+\sqrt{65})/10,\ 0,\ (25+\sqrt{65})/20]$ on C; $[2,\ (\sqrt{65}-7)/2,\ 2]$ on B and F; $[(13+\sqrt{65})/26,\ 2,\ (3\sqrt{65}-13)/52]$ on D; $[2/3,\ (19-\sqrt{65})/6,\ 0]$ on E; and p; figure 3c.

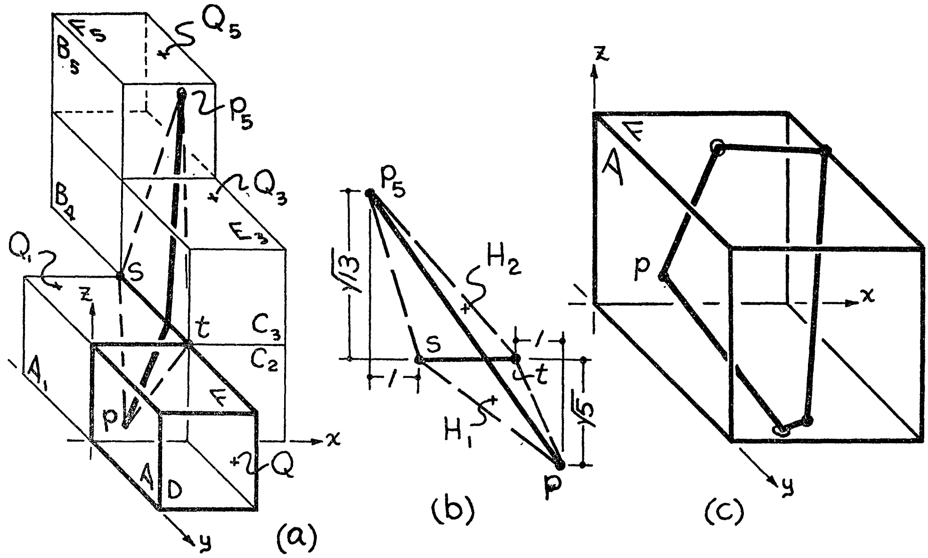


FIG. 3

Problem 1c, *Solution.* Reflecting Q consecutively through faces C-D-F-E-B, figure 4a; we see that the shortest path will lie in planes H_3 and H_4 determined by points p, s, $u(0, -2, 2)$ and p_5, $v(2, -2, 4)$, $w(2, -4, 4)$, respectively.

This observation seems to be all we can readily determine geometrically. Thus we employ calculus to conclude the solution.

Let X be a point x units from point u on line $\overline{us}$. Referring to figure 4b, it is evident that the shortest path through X connecting p and p_5 is

$$L(x) = \sqrt{10 + x^2} + \sqrt{14 - 4x + x^2 + 2\sqrt{5(8 - 4x + x^2)}}, \qquad 0 \leqq x \leqq 2.$$

Minimizing $L(x)$, we find that $x=1.21$ correct to two decimals and this path when reflected back into cube Q connects the points p; (.40, 0, 1.33) on face C; (1.21, 2, 2) on faces D and F; (2, 1.51, 0) on faces E and B; and p; figure 4c. Although the solution of problem 1c eventually employs calculus involving a function of one variable, the geometric attack certainly simplified the virtually impossible problem of minimizing a function of ten variables arising from a purely analytic viewpoint.

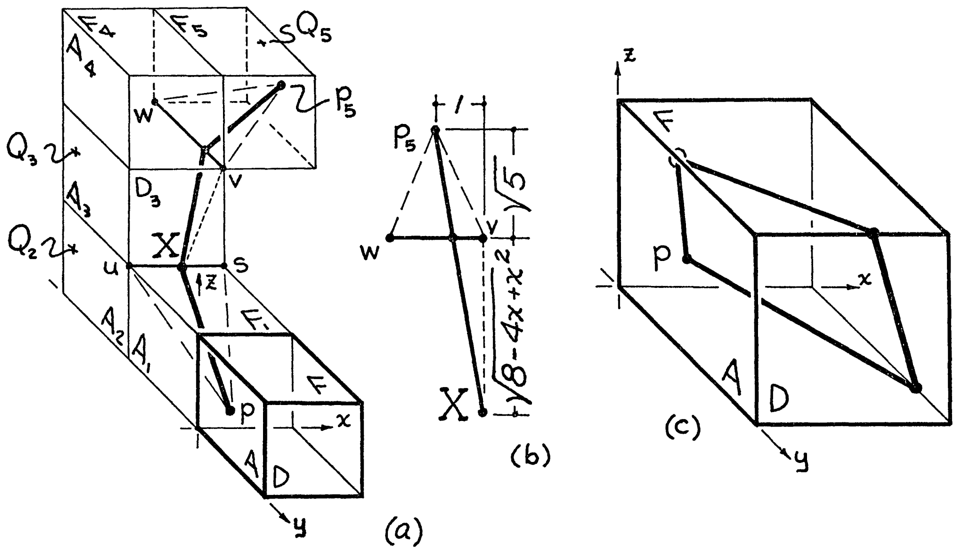


FIG. 4

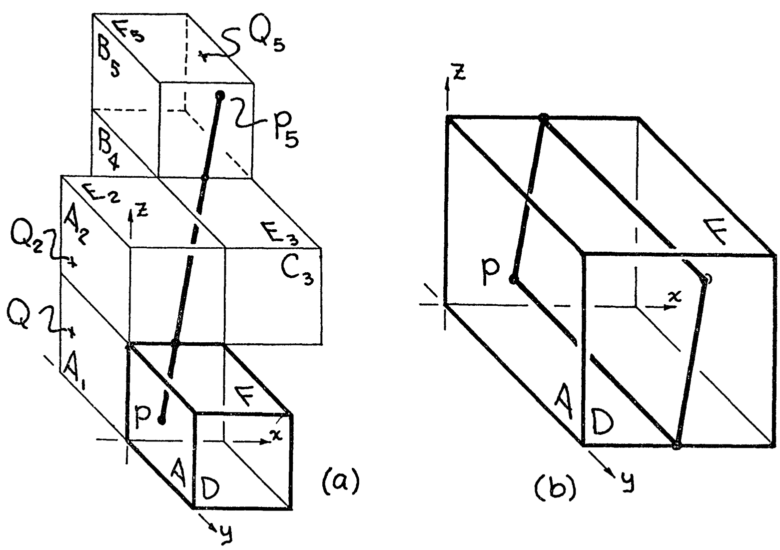


FIG. 5

Problem 2, *Solution.* From the preceding discussion, it becomes apparent that any sequence of reflections through each of the five faces will put p_5 at one of the four points $(4, 5, 5)$, $(4, 5, -3)$, $(4, -3, 5)$ or $(4, -3, -3)$, all of which are $4\sqrt{3}$ units from the point p. Thus if some sequence of reflections produces a polyhedron such that p and p_5 can be connected by a straight line lying within the

region we would certainly have the shortest path. Indeed, the sequence C-F-B-D-E produces such a polyhedron, figure 5a, and the desired path connects points: p; (1, 0, 2) on faces C and F; (2, 1, 1) on face B; (1, 2, 0) on faces D and E; and p; figure 5b.

It should be noted that this solution is not unique. In fact, there are four possible straight line solutions giving two distinct minimal paths. Furthermore, a more irregular polyhedron could very well have the final image of point p at various distances from p. Thus, in general, one must investigate all possible paths to determine the shortest path.

In conclusion, we point out that although this paper was restricted to closed paths within cubes, the method of attack can be readily employed to simplify a variety of problems involving paths of minimal length within quite general polyhedrons.

References

1. A. V. Baez, Fermat's principle and certain minimum problems, this MONTHLY, 55 (1948) 316.
2. A. A. Bennett, On the treatment of maxima and minima in calculus texts, this MONTHLY, 31 (1924) 293–294.
3. R. Garver, The solution of problems in maxima and minima by algebra, this MONTHLY, 42 (1935) 435–437.
4. R. A. Johnson, Relating to a problem in minima, this MONTHLY, 24 (1917) 243–244.
5. N. J. Lennes, Note on maxima and minima by algebraic methods, this MONTHLY, 17 (1910) 9–10.

MATRICES OVER A FINITE FIELD

S. D. FISHER, University of Wisconsin and M. N. ALEXANDER, Cornell University

This paper is concerned with matrices over the field of q elements, where, of course, $q=p^N$, p a prime and N a positive integer.

Theorems 1–5 are computations of the orders of certain subsets of these matrices while Theorem 6 describes some of the properties of the matrices themselves. Theorems 1 and 2 are well-known results but the proof of Theorem 2 appears to be new.

Some of the proofs make use of orbits and stabilizers. We give the basic definitions and theorem below; additional details may be found in Chevalley (2).

Let H be a group, S a set, and G_S the group of one-to-one functions from S onto S. H acts on S if there is a homomorphism, f, from H into G_S. If this is the case, the orbit of a fixed s in S is the set of all elements of S of the form $f(h)(s)$, for h in H. The stabilizer of a fixed s in S is the set of all h in H such that $f(h)(s)=s$. In the case when H is finite, we have the following relationship: the order of the orbit of s is the quotient of the order of H and the order of the stabilizer of s. The proof of this theorem is obtained by noting that the function which associates with each left coset, hH_s, of the stabilizer of s in H the element $f(h)(s)$, in the orbit of s, is both one-to-one and onto.

We denote by F_{nm} the vector space of $n\times m$ matrices and by G_n the group of $n\times n$ nonsingular matrices over this field.

THEOREM 1. *The order of* G_n *is* $\prod_{j=0}^{n-1}(q^n-q^j)$.

Proof. Given s $n\times 1$ vectors, there are q^s linear combinations of them and hence q^n-q^s $n\times 1$ vectors linearly independent of them. ∎

THEOREM 2. *The number of* $n\times m$ *matrices of rank* k *is*

$$\frac{\prod_{j=0}^{k-1}(q^n-q^j)\prod_{j=0}^{k-1}(q^m-q^j)}{\prod_{j=0}^{k-1}(q^k-q^j)}.$$

Proof. A matrix A has rank k if and only if A is equivalent to the $n\times m$ matrix

$$H=\begin{bmatrix} I_k & 0\\ 0 & 0\end{bmatrix}$$

where I_k is the $k\times k$ identity.

Let $G_n\oplus G_m$, the direct sum of G_n and G_m, act on M, the set of $n\times m$ matrices of rank k, by $(g_n, g_m)A\to g_nAg_m^{-1}$ for $g_n\in G_n$, $g_m\in G_m$, and $A\in M$. The orbit of H under this action is all of M. Hence, to find the order of M we need only find the order of S, the stabilizer of H.

Suppose $U=(u_{ij})$ and $V=(v_{ij})$ and $(U, V)\in S$. Then $UH=HV$ and consequently we find

$$U=\begin{bmatrix} W & \mathrm{I}\\ 0 & \mathrm{II}\end{bmatrix} \quad\text{and}\quad V=\begin{bmatrix} W & 0\\ \mathrm{III} & \mathrm{IV}\end{bmatrix}$$

where W is $k\times k$ and W, I, II, III, IV are arbitrary, provided only that U and V are nonsingular. By Theorem 1, W may be formed in $\prod_{j=0}^{k-1}(q^k-q^j)$ ways.

Once W is formed it follows that the remainder of U may be formed in $\prod_{j=k}^{n-1}(q^n-q^j)$ ways so that U is nonsingular. Similarly, the remainder of V may be formed in $\prod_{j=k}^{m-1}(q^m-q^j)$ ways making V nonsingular. Therefore, the order of S is

$$\prod_{j=0}^{k-1}(q^k-q^j)\prod_{j=k}^{n-1}(q^n-q^j)\prod_{j=k}^{m-1}(q^m-q^j)$$

and the conclusion follows from Theorem 1. ∎

The proof of the next theorem is immediate and is therefore omitted.

THEOREM 3. *Let* v *be a nonzero fixed element of* F_{m1}. *Let* $S_0=\{T\in F_{nm}\mid Tv=0\}$. *Then* F_{nm}/S_0 *is isomorphic to* F_{n1} *as groups; hence, if* $w\in F_{n1}$, *then the number of* $T\in F_{nm}$ *such that* $Tv=w$ *is* $q^{n(m-1)}$.

THEOREM 4. *Let v be a fixed nonzero element of F_{n1}. The number of $T \in G_n$ such that $Tv = v$ is $\prod_{k=1}^{n-1} (q^n - q^k)$; hence, each nonzero v in F_{n1} induces a decomposition of G_n into $q^n - 1$ disjoint subsets, $S_w = \{T \mid Tv = w\}$, each of order*

$$\prod_{k=1}^{n-1} (q^n - q^k),$$

where w ranges over the nonzero elements of F_{n1}.

Proof. Let G_n act on the nonzero elements of F_{n1} by $T \to Tv$. Then the orbit of v is all of these elements and therefore the order of the stabilizer of v is the order of G_n divided by the order of F_{n1}. ∎

Let T be an $n \times m$ matrix. We say that the ith row of T has character k if $\sum_{j=1}^{m} t_{ij} = k$. If each row has character k, then T is a matrix of character k. Note that T has character k if and only if $T\theta = k\theta$, where θ is the column vector of all 1's.

THEOREM 5. (1) *The number of $n \times m$ matrices with n_k rows of character k, where $\sum_{k=0}^{q-1} n_k = n$, is $[n!/n_0! \cdots n_{q-1}!]q^{n(m-1)}$.*

(2) *The number of nonsingular matrices with n_k rows of character k, where $\sum_{k=0}^{q-1} n_k = n$, is $[n!/n_0! \cdots n_{q-1}!] \prod_{j=1}^{n-1} (q^n - q^j)$.*

Proof. There are $[n!/n_0! \cdots n_{q-1}!]$ elements of F_{n1} with exactly n_k k's. Conclusions (1) and (2) now follow from Theorems 3 and 4, respectively. ∎

THEOREM 6. (1) *If T is nonsingular and has character k, then T^{-1} has character k^{-1} (note that it is impossible that $k = 0$).*

(2) *If A has character nk and B has character mk, then $A + B$ has character $(n+m)k$.*

Hence, the nonsingular matrices of character k^t, t an integer, form a subgroup of G_n which by Theorem 5 has order

$$\epsilon_k \prod_{j=1}^{n-1} (q^n - q^j),$$

where ϵ_k is the number of distinct powers of k (mod q); *the matrices of character sk, s an integer, form a subgroup of F_{nm} of order $\eta_k q^{n(m-1)}$ where η_k is the number of distinct multiples of k* (mod q).

Proof. The proof is immediate by noting the effect of multiplying θ by the prescribed matrices. ∎

This paper was written while the authors were employed for the summer at the Office of Research and Development, United States Patent Office.

References

1. G. Birkhoff and S. MacLane, A Survey of Modern Algebra, rev. ed., Macmillan, New York, 1953.
2. C. Chevalley, Fundamental Concepts of Algebra, Academic Press, New York, 1957.
3. L. E. Dickson, Linear Groups, Teubner, Leipzig, 1901.

THE OUTER AUTOMORPHISMS OF S_6

P. J. LORIMER, University of Canterbury, Christchurch, New Zealand

1. It is a well-known fact of group theory, that the symmetric group S_n is complete if $n \geqq 3$, $n \neq 6$. See, for example, [1] p. 92. The group S_6 has 720 outer automorphisms which form a single coset of the group of inner automorphisms. It is the purpose of this paper to give these automorphisms explicitly in terms of a set of generators of S_6.

2. According to [2] p. 64, $R_1=(12)$, $R_2=(23)$, $R_3=(34)$, $R_4=(45)$, and $R_5=(56)$ generate S_6 and we have the following defining relations:

$$\begin{array}{llll} \text{D1.} & R_i^2 = 1 & & 1 \leqq i \leqq 5 \\ \text{D2.} & (R_iR_{i+1})^3 = 1 & & 1 \leqq i \leqq 4 \\ \text{D3.} & (R_iR_k)^2 = 1 & & i \leqq k-2. \end{array}$$

3. From the usual proof of the completeness of S_n, $n \geqq 4$ ([1] p. 92) it is clear that any outer automorphism of S_6 must map a transposition (ab) onto a product of three disjoint transpositions. We denote the set of these products $(cd)(ef)(gh)(c, \cdots, h$ all different) by C_3, and the set of all transpositions by C_1. Thus, whereas an inner automorphism maps C_1 onto itself, an outer automorphism maps C_1 onto C_3.

We first exhibit a particular automorphism of this type.

PROPOSITION 1.

$$\begin{aligned} T_1 &= (12)(34)(56) \\ T_2 &= (35)(61)(24) \\ T_3 &= (12)(36)(45) \\ T_4 &= (34)(61)(25) \\ T_5 &= (12)(35)(46) \end{aligned}$$

are a set of generators of S_6.

Proof. It is easily shown that

$$\begin{aligned} (12) &= [1\ 3\ 5] \\ (23) &= [24](45) \\ (34) &= [1\ 3\ 5\ 4\ 5\ 3\ 4] \\ (45) &= [1\ 5\ 2\ 3\ 1\ 2\ 5](34)[2\ 3\ 1\ 2\ 5] \\ (56) &= [1\ 4\ 5\ 3\ 4] \end{aligned}$$

where [1 3 5] stands for the permutation $T_1T_3T_5$; etc.

PROPOSITION 2. $T_1 \cdots T_6$ *satisfy the relations*:

$$
\begin{array}{lll}
\text{D1.} & T_i^2 = 1 & 1 \leqq i \leqq 5 \\
\text{D2.} & (T_iT_{i+1})^3 = 1 & 1 \leqq i \leqq 4 \\
\text{D3.} & (T_iT_k)^2 = 1 & i \leqq k - 2.
\end{array}
$$

This proposition is an immediate consequence of:

LEMMA 1. *If α and β are permutations from C_3, then $(\alpha\beta)^2 = 1$ if and only if α and β have a transposition in common. Otherwise $(\alpha\beta)^3 = 1$.*

Proof. If $\alpha = (ab)(cd)(ef)$ and $\beta = (ab)(ce)(df)$ then $\alpha\beta = (cf)(de)$. If $\alpha = (ab)(cd)(ef)$ and $\beta = (ac)(be)(df)$ then $\alpha\beta = (ade)(bfc)$.

It is clear that all cases come under one of these two categories.

From Propositions 1 and 2 it can now be seen that the mapping $R_i \to T_i$ $i = 1 \cdots 5$, induces an outer automorphism of S_6.

4. We now show that there are exactly 720 ways in which the permutations $T_1 \cdots T_6$ can be chosen so that they generate S_6 and satisfy D1, 2, 3. We recall that $T_1 \cdots T_6$ must be permutations of C_3.

(i) We first choose T_1. This can be done arbitrarily from C_3 and hence may be chosen in 15 ways.

(ii) Suppose T_1 is chosen. By D2 we must have $(T_1T_2)^3 = 1$ and hence, by Lemma 1, T_2 and T_1 can have no transposition in common. T_2 can be chosen in 8 ways.

(iii) By Lemma 1 and D2, D3, it can be seen that T_3 must have a transposition in common with T_1 but no transposition in common with T_2. T_3 can be chosen in 3 ways.

(iv) T_4 must have a transposition in common with T_1 and T_2, but no transposition in common with T_3. T_4 can be chosen in 2 ways.

(v) T_5 is determined uniquely by $T_1 \cdots T_4$.

Thus $T_1 \cdots T_5$ can be chosen in 720 ways, each satisfying D1, 2, 3; and it is clear that each of these choices, together with the mapping $R_i \to T_i$ determines an outer automorphism of S_6.

5. The group of inner automorphisms maps C_1 onto itself. We can choose as generators of S_6, the transpositions $U_1 = (ab)$, $U_2 = (bc)$, $U_3 = (cd)$, $U_4 = (de)$ and $U_5 = (ef)$, where

$$
\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ a & b & c & d & e & f \end{pmatrix}
$$

is any permutation of S_6. Then the mappings $R_i \to U_i$ $(i = 1 \cdots 5)$ induce the group of inner automorphisms of S_6. There are clearly 720.

References

1. A. G. Kurosh, The Theory of Groups, vol. 1. Chelsea, New York, 1960.

2. H. S. M. Coxeter and W. O. J. Moser, Generators and Relations for Discrete Groups, Springer-Verlag, Berlin, 1957.

PROJECTION OF A QUADRANGLE INTO A PARALLELOGRAM

W. H. RICHARDSON, Wichita State University

In [1], problem 2 on page 9 reads "Show that any quadrangle may be projected into a parallelogram." The solution of this problem usually goes as follows: Let $PQRS$ be the given quadrangle with exterior diagonal points T and U. Let O be any point not in the plane of $PQRS$ and join the points P, Q, R, S, T, U with O. See figure 1. Now if the set of lines OP, OQ, OR, OS are cut by a plane α parallel to the plane OTU we get points P' on OP, Q' on OQ, R' on OR and S' on OS, which when joined form a parallelogram.

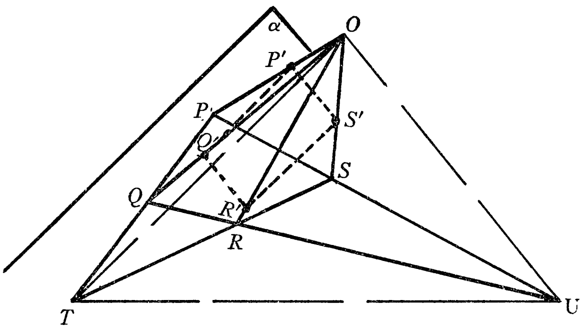


FIG. 1

The purpose of this note is to investigate the case in which O is in the plane of $PQRS$. More precisely, given a quadrangle $PQRS$ and a point O in its plane, construct a line X_1X_2 such that there is a parallelogram $P'Q'R'S'$ which is perspective with $PQRS$ from the point O and also from the line X_1X_2. See figure 2.

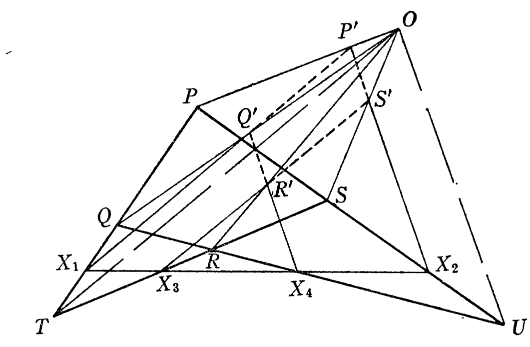


FIG. 2

If O is in the plane of $PQRS$, the construction is as follows. Join O with the vertices P, Q, R, S and the diagonal points T and U. On the line OP pick a point P' and construct lines $P'X_1$ and $P'X_2$ parallel to OT and OU, respectively. $P'X_1$ meets OQ in a point Q' and $P'X_2$ meets OS in a point S'. Join the points

X_1 and X_2. The line X_1X_2 meets ST at X_3 and QU at X_4. Join the points X_3 and S' which meet OR at R'. Join $Q'X_4$. This line intersects OR at R' and the figure $P'Q'R'S'$ is the required parallelogram.

Proof. Triangles $OP'Q'$ and UX_2X_4 are perspective from the line PQ. By Desargues' Theorem, these triangles are perspective from some point. By construction, $P'X_2$ is parallel to OU. Therefore, the center of perspectivity is an ideal point, and thus $Q'X_4$ is parallel to $P'X_2$. Since triangles $OP'S'$ and TX_1X_3 are perspective from the line PS, we see, by a similar argument, that $S'X_3$ is parallel to $P'X_1$. All that remains to prove is that $Q'X_4$, $S'X_3$ and OR are concurrent at R'.

Suppose $S'X_3$ meets OR at R'. Triangles $OR'S'$ and UX_4X_2 are perspective from the line RS, and therefore they must be perspective from a point. Line $S'X_2$ was constructed parallel to OU; therefore $R'X_4$ is parallel to OU. On the other hand, it has already been established that $Q'X_4$ is parallel to OU. Since one and only one line can be drawn parallel to a given line through a given point, R' is on line $Q'X_4$. Therefore, lines OR, $Q'X_4$, $S'X_3$ are concurrent on R'.

It is worth noting that if O is chosen to be on the circle with diameter TU, the parallelogram $P'Q'R'S'$ will be a rectangle. An interesting problem is to determine the positions of O, on this circle (if there are any), for which $P'Q'R'S'$ is a square.

Reference

1. T. F. Holgate, Projective Pure Geometry, Macmillan, New York, 1930.

A CHARACTERIZATION OF A PARTIAL-BOUNDARY SET FUNCTION

J. E. Joseph, Howard University

Frank Reese Harvey [1] and Hyman Gabai [2] present derived set, exterior and boundary operator axioms which characterize a topological space. In this paper the term partial boundary of a set will mean the intersection of the set and the closure of the complement of the set, and a topological space will be characterized by partial-boundary operator axioms. The following example shows that the boundary and partial boundary of a set do not coincide in general.

Example. Consider the reals with the usual topology and let $A=[0, 1)$. Then the boundary of A is $\{0, 1\}$ and the partial boundary of A is $\{0\}$.

The reader can easily verify the following relationships between the boundary and partial boundary of a set.

I. The boundary of a set is the union of the partial boundaries of the set and its complement.

II. The boundary and partial boundary of a set are equal if and only if the set is closed.

DEFINITION. *A partial boundary operator on a nonempty set X is a function, ∂, which maps the subsets of X into the subsets of X and which satisfies the following statements*:

(PB1) $\partial X=\phi$.
(PB2) $\partial[c\partial(cA)]\subseteq A$ *for each* $A\subseteq X$ (cA *denotes the complement of* A *in* X).
(PB3) $\partial(A\cap B)=(A\cap\partial B)\cup(B\cap\partial A)$ *for each* $A\subseteq X$ *and* $B\subseteq X$.

The following shows that boundary satisfies PB1–PB3 in a space for each $A\subseteq X$ and $B\subseteq X$. Let $A^*=$partial boundary of $A=A\cap\overline{cA}$.

$$\text{(PB1)}\quad X^*=X\cap\overline{cX}=X\cap\bar{\phi}=\phi.$$

$$\begin{aligned}\text{(PB2)}\quad [c(cA)^*]^* &= c(cA)^*\cap\overline{(cA)^*}\\ &= c(cA\cap\overline{A})\cap\overline{(cA\cap\overline{A})}\\ &\subseteq (A\cup c\overline{A})\cap(\overline{cA}\cap\overline{\overline{A}})=A\cap\overline{cA}\subseteq A.\end{aligned}$$

$$\begin{aligned}\text{(PB3)}\quad (A\cap B)^* &= (A\cap B)\cap\overline{c(A\cap B)}=(A\cap B)\cap\overline{(cA\cup cB)}\\ &= (A\cap B)\cap(\overline{cA}\cup\overline{cB})=(A\cap B\cap\overline{cB})\cup(A\cap\overline{cA}\cap B)\\ &= (A\cap B^*)\cup(B\cap A^*).\end{aligned}$$

The following lemma will be used in proving the main result.

LEMMA. *If ∂ is a partial boundary operator on X, and $A\subseteq X$, then $\partial A\subseteq A$.*

Proof. If in PB3, $A=B$, then $\partial(A\cap A)=A\cap\partial A$ implies $\partial A\subseteq A$.

THEOREM. *If ∂ is a partial-boundary operator on a nonempty set X, there exists a unique topology, O, on X such that for each $A\subseteq X$, $\partial(A)=A^*$.*

Proof. Define a map k of the subsets of X into the subsets of X by $k(A)=A\cup\partial(cA)$ for each $A\subseteq X$. The following shows that k satisfies Kuratowski's axioms:

$$\text{(K1)}\quad k(\varnothing)=\varnothing\cup\partial X=\varnothing\cup\varnothing=\varnothing.$$

For each A, $B\subseteq X$,

$$\text{(K2)}\quad k(A)\supseteq A,$$

$$\begin{aligned}\text{(K3)}\quad k(k(A)) &= A\cup\partial(cA)\cup\partial(cA\cap c\partial(cA))\\ &= A\cup\partial(cA)\cup[cA\cap\partial(c\partial(cA))]\cup[\partial(cA)\cap c\partial(cA)]\\ &= k(A),\end{aligned}$$

$$\begin{aligned}\text{(K4)}\quad k(A\cup B) &= (A\cup B)\cup\partial[c(A\cup B)]=(A\cup B)\cup\partial(cA\cap cB)\\ &= (A\cup B)\cup[cA\cap\partial(cB)]\cup[cB\cap\partial(cA)]\\ &= (A\cup B)\cup[cA\cap cB\cap\partial(cB)]\cup[cB\cap cA\cap\partial(cA)],\end{aligned}$$

by the lemma.

$$\begin{aligned} k(A \cup B) &= (A \cup B) \cup [cA \cap cB \cap (\partial(cA) \cup \partial(cB))] \\ &= A \cup B \cup [\partial(cA) \cup \partial(cB)] \\ &= [A \cup \partial(cA)] \cup [B \cup \partial(cB)] = k(A) \cup k(B). \end{aligned}$$

k induces a unique topology, 0, on X such that for each $A \subseteq X$, $k(A) = \bar{A}$. In $(X, 0)$, $A^* = A \cap \overline{cA} = A \cap k(cA) = A \cap [cA \cup \partial A] = A \cap \partial A = \partial A$ for each $A \subseteq X$.

The following proof of uniqueness of k is due to the referee. Suppose that k' is another closure function on X such that $\partial A = A \cap k'(cA)$ for every $A \subseteq X$ and $k' \neq k$. Then for some A, $k(cA) \neq k'(cA)$. Then, there exists a point, say x, in $k(cA)$ but not in $k'(cA)$ (or vice versa). This x must be in A. Hence, $x \in A \cap k(cA)$, $x \notin A \cap k'(cA)$. But $A \cap k(cA) = A \cap k'(cA)$ which is a contradiction.

The author is indebted to the referee for his suggestions. He pointed out the similarity of PB3 to the formula for the derivative of a product in differential calculus and to the behavior of the boundary operator in algebraic topology.

References

1. F. R. Harvey, The derived set operator, this MONTHLY, 70 (1963) 1085–1086.
2. H. Gabai, The exterior operator and boundary operator, this MONTHLY, 71 (1964) 1029–1031.
3. N. Bourbaki, Topologie Générale, Éléments de Mathématique, Livre III, Hermann, Paris, 1960.

MODULES OVER COMMUTATIVE RINGS

NEAL H. McCOY, Smith College

Let R be an arbitrary commutative ring with unity and M a unitary R-module. Elementary proofs have been given in two recent notes ([1], [2]) that *if M has a basis of n elements, then every basis of M contains exactly n elements*. The following is still another comment on the proof of this result.

In [3] there is defined the concept of rank of a matrix over R (coinciding with the classical definition in case R is an integral domain) and it is proved that a system of linear homogeneous equations over R has a nontrivial solution in R if and only if the rank of the matrix of coefficients is less than the number of unknowns. In particular, *a system of linear homogeneous equations over R with fewer equations than unknowns always has a nontrivial solution in R*. Precisely as in the case of a vector space over a field, this result shows that if M is generated by n elements, any $n+1$ elements of M are dependent. The stated theorem then follows immediately.

References

1. G. A. Kandall, The rank of an R-module, this MONTHLY, 70 (1963) 653–655.
2. W. G. Leavitt, Modules over commutative rings, this MONTHLY, 71 (1964) 1112–1113.
3. N. H. McCoy, Rings and Ideals, Carus Mathematical Monograph No. 8, MAA, 1948.

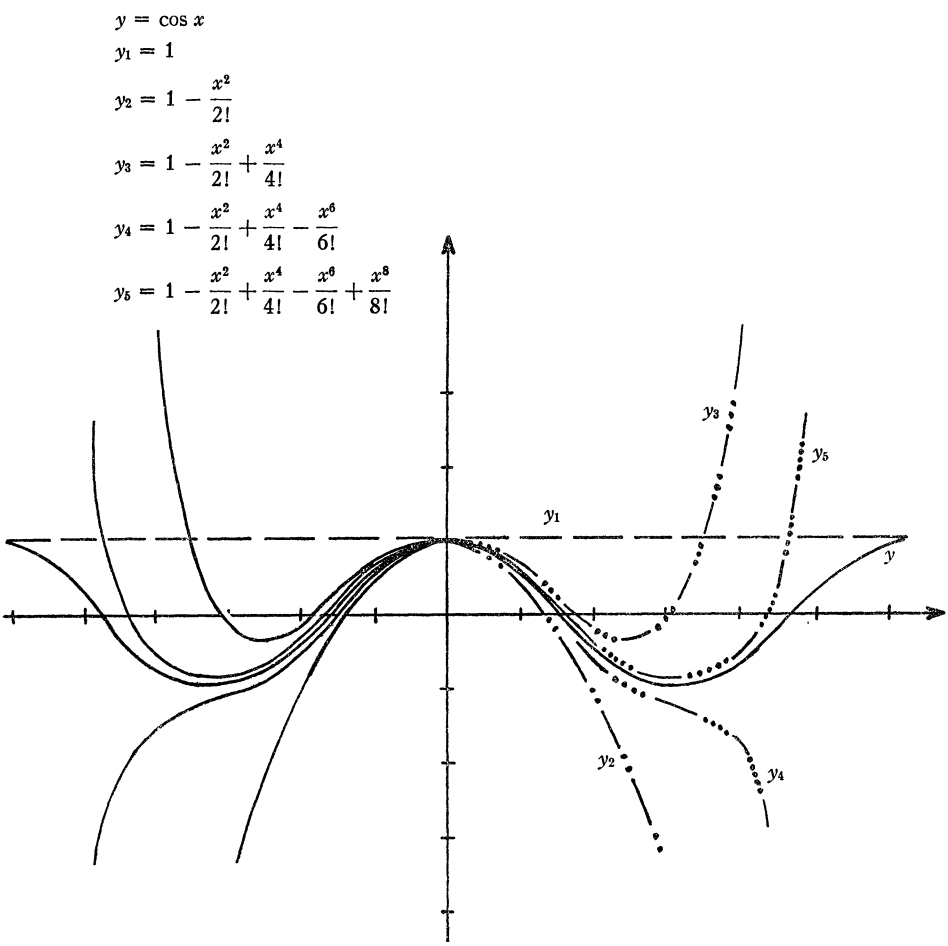
GEOMETRIC INTERPRETATIONS OF POLYNOMIAL APPROXIMATIONS OF THE COSINE FUNCTION

E. R. HEINEMAN, Texas Technological College

The polynomial approximations obtained from the power series expansion

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \tag{1}$$

converge to the cosine function as shown.



It is interesting to notice that if $y_i(x)$ is the ith polynomial approximation obtained from (1), then

$$\frac{d^2y_i}{dx^2} = - y_{i-1},$$

which shows that the inflection points of the graph of y_i have the same abscissas as the x-intercepts of y_{i-1}.

In one sense, it is clear that the successive approximations become "better." The student, however, must be warned against the pseudo-truth: "Every time you add another term, the approximation gets better." The horizontal line $y_1 = 1$ is a shabby approximation for the cosine curve, but it is never in error by more than 2, which is more than can be said about the later approximations.

The addition of another term to y_2 gives y_3, which is a better approximation for $\cos x$, but *only if* $|x| < \sqrt{12+\sqrt{96}} \approx 4.67$. For $x > 2\sqrt{3+\sqrt{6}}$, the three-term polynomial y_3 is a worse approximation than the two-term polynomial y_2.

The addition of two terms to y_2 produces y_4, which is a better approximation for $\cos x$, but only if $|x| < \sqrt{30}$. The graph of y_4 intersects that of y_2 at $(\pm\sqrt{30}, -14)$. For $|x| > \sqrt{30}$ it can be shown that $y_4 < y_2$, the graph of y_4 lies below that of y_2, and y_4 is a worse approximation than y_2.

BINOMIAL DIVISIBILITY BY PRIME POWERS

MARLOW SHOLANDER, Western Reserve University

We give a slight extension of remarks on divisibility in [1] and [2]. By expanding $v^p = (kp^r - u)^p$ and examining terms, we readily establish the following

THEOREM. *If p is an odd prime not a factor of u, and $r \geqq 1$, then $u+v \equiv 0 \pmod{p^r}$ implies $u^p+v^p \equiv 0 \pmod{p^{r+1}}$. Further, $u+v \not\equiv 0 \pmod{p^{r+1}}$ implies $u^p+v^p \not\equiv 0 \pmod{p^{r+2}}$.*

COROLLARY.

$$u + v \equiv 0 \pmod{p^r} \Leftrightarrow u^p + v^p \equiv 0 \pmod{p^{r+1}} \Leftrightarrow \cdots$$
$$\Leftrightarrow u^{p^k} + v^{p^k} \equiv 0 \pmod{p^{r+k}}.$$

Similar implications, in one direction at least, also hold for any odd modulus.

References

1. L. M. Court, A divisibility theorem, this MONTHLY, 71 (1964) 300–301.
2. Lincoln LaPaz, Chains of factorable binomials, this MONTHLY, 72 (1965) 58–59.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and
JOHN A. BROWN, University of Delaware

Address all correspondence to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

ALGEBRAIC GEOMETRY FOR HIGH SCHOOL PUPILS

S. S. WILLOUGHBY, New York University

Geometry as a school subject has been under attack since J. J. Sylvester wished publicly for the interment of Euclidean Geometry "'deeper than e'er plummet sounded' out of the schoolboy's reach." Dieudonné's more recent pronouncement that "Euclid must go" was simply a stronger version of what some mathematicians had been saying for many years.

Assuming Euclid *does* go, what shall we use to replace him? There has been considerable pressure, of late, to replace Euclid with a more algebraic kind of geometry, and even some tendency to replace all references to geometry by topics which most teachers do not consider to be geometry at all [1]. In considering the matter of what materials ought to be taught in the time traditionally allotted to Euclidean geometry, there are two distinct questions to be answered. The first of these is: "*Can* the material in question be taught to high school pupils?" The second is: "*Ought* the material to be taught to high school pupils?" Presumably, the set of materials for which the second question is answered affirmatively is a proper subset of the set for which the first question is answered affirmatively.

An experiment which was carried out in Greenwich (Connecticut) High School during the 1958–59 academic year, and a subsequent follow-up study may throw some light on both the question of what can and what ought to be taught to tenth grade mathematics pupils.

The material used for the course was highly algebraic and began with the geometry (or algebra) of the line. An affine geometry was developed, and many theorems were proved before Euclidean restrictions were placed on the geometry. There were no pictures in the text although the teacher provided many pictures in the classroom. It seemed to the teacher that geometric intuition played a lesser role in the pupil's understanding of the course than algebraic intuition. All of the usual topics of a traditional geometry course were covered in the course, though some received less emphasis than is customary, needless to say. There was also a substantial introduction to trigonometry.

The experimental class consisted of nine boys and fourteen girls with Otis I.Q. scores ranging from 106 to 127. All members of the class had been in a ninth grade course which was based on material prepared by the University of Illinois Committee on School Mathematics which covered much the same material as the usual ninth grade course with somewhat more emphasis on understanding

and thought by the pupil rather than manipulation of symbols and formulas. It also included a section on graphs and ordered pairs which fitted into the geometry course nicely.

Soon after the course began, it became clear that the algebraic background of the pupils was not sufficient. In order to achieve real understanding of the proofs it was often necessary to detour first into topics of algebra which the text assumed. Although these excursions were undoubtedly of value to the pupils, they tended to interfere with the continuity of thought and to delay progress with the geometry. A group of students with a stronger algebraic background would have been able to concentrate more energy on the geometry. Considering the recent trend toward teaching some algebra to accelerated groups before the ninth grade, it will probably not be difficult to obtain such classes in the future.

During most of the course the text was followed quite closely, and the pupils quickly learned to appreciate the beauty of the proofs and the need for them in a mathematical structure. Near the end of the course, it was necessary to move more rapidly in order to finish, and plausible inductive reasoning often took the place of the deductive proofs of the last four chapters.

At the time, the enthusiasm of the pupils was great. Undoubtedly, part of the reason for this was the uniqueness of their situation and the novelty of the entire experiment. The opportunity to see and talk with the author of the text materials clearly impressed the pupils greatly. Also, the knowledge that they were the first class ever to see some of the material was thrilling to many of the group. Much of the enthusiasm, however, was obviously generated by an interest in the material and the challenge which it offered. There was a clear desire on the part of the pupils to learn more of this fascinating subject—mathematics. In this respect, it is interesting to note that every member of the class chose to continue to take mathematics the next year.

Two good students, who were taking standard geometry courses, asked for permission to change into this class (having heard good comments about it from members of the class). Each managed to study enough on his own, with help after school, to pass tests on the material the class had studied, and were thus allowed to change. One joined the group after six weeks and the other after twelve weeks. At the end of the year, both assured the teacher that they thought if was easier because it dealt more with big ideas than with details.

At the end of the year the students took the Cooperative Plane Geometry Test. Scores on this test ranged from 52 to 76. (The National mean is 50 and the standard deviation is 10.) These scores are particularly surprising since much of the test involved a knowledge of, or an ability to figure out, proofs from a usual plane geometry course. Presumably, a test on the material of this course for a group from a regular course would have been useless, since they could hardly be expected to know enough of the material to write anything intelligible about it.

Follow-up study. It has become commonplace for educational experiments to demonstrate that remarkable things can be done in teaching young pupils concepts which seem difficult to adults. The results above supply further corroboration. A more interesting question is whether or not the material ought to be taught to the pupils.

At least part of the answer to the question of desirability should come from the pupils themselves. After several years (six in this case, giving individuals a chance to complete substantial college work), what do they think of the course, how do they think it has changed their lives, do they believe it was a more worthwhile experience than they would have received otherwise?

Of course, there are problems in collecting this sort of data. How does a pupil know what his reactions might be to another course? Will he blame a course unfairly for his subsequent failures or credit it unfairly for his success? Will he forget significant facts? Will the teaching or the experimental nature of the course affect unduly the judgment of the pupil? All of these difficulties exist and must be considered when evaluating the following information. Even considering these factors, however, we can draw some interesting tentative conclusions.

A questionnaire was sent to all participants in the course asking the following four questions:

1. What courses in mathematics have you taken since tenth grade geometry? What were your grades in these courses (if you remember)?
2. What courses (such as physics, engineering, etc.) which are closely related to mathematics have you taken since tenth grade? What were your grades in these courses (if you remember)?
3. In your opinion, what effects, good or bad, resulted from your taking of the special geometry course rather than the usual course? (That is, did it affect your desire to take other courses, did it affect your performance in other courses, did it affect your life and activities outside of school, etc.?)
4. Make any further comments about the course which you think might be of interest. Use the other side of this sheet if necessary.

Unfortunately, some of the pupils had moved or did not respond for other reasons. Four of the 25 pupils could not be located and their questionnaires were returned by the post office, unopened. Two of the pupils simply did not respond for reasons unknown to the author.

Of the 19 who responded, one was violently opposed to the course. It is interesting to note that she is presently married to one of her college mathematics instructors. Three others felt that the course did them more harm than good, but were not as violent as the young lady mentioned above. Five of the respondents were either neutral or rather vague. Seven generally approved of the course and thought it had helped them substantially. Three were as violent in their praise of the course as the first young lady was in her condemnation.

There are many interesting questions one can ask regarding the kinds of reasons the students may have had for reacting as they did. For example, did the grade in the course affect opinions? It happens that the two people with the two lowest grades thought the course had hurt them more than helped, but future success in the course had been predicted from previous grades and teachers opinions, and there was no obvious correlation between opinions and how the pupils had done relative to expectations.

Does I.Q. affect the respondent's opinion? Apparently it does, but in a rather unexpected way. The people who expressed strong opinions (either way) had consistently high I.Q.'s (average above 125) while the people who were moderately favorable towards the course had the lowest I.Q.'s—an average of almost ten points lower. This fact may be entirely independent of the course under discussion.

Some of the comments made by the respondents were enlightening. The most common complaint was that the course did not help on the College Entrance Examinations; another rather common comment was that the course did help in college mathematics courses. Only two people said the course helped on the College Entrance Examinations while two others said that the regular geometry course would have been more useful in later courses in science or mathematics.

There was some tendency to comment on the actual teaching of the course rather than the content, but for the most part it was the content that was discussed. Several of the objections derived from the fact that the text material used was still in a developmental stage and there really were not enough exercises; thus the pupils were not given as much homework practice as they might normally expect. On the other hand, several commented on the emphasis on thinking and concepts rather than on "busy work"—this also may have derived from the lack of available practice material. Several people suggested that the course would be more worthwhile if it had come later in their formal education and been integrated more closely with the other mathematics courses they took. This undoubtedly is true.

In evaluating the program, we must take into consideration the probable Hawthorne Effect, and the likely desire of the individuals to say something that they thought would please their former teacher, as well as the apparent successes of the program in terms of standardized tests and attitudes. Taking all of these things into consideration, and weighing them as best I can, it is my opinion that this was a very interesting experience for both pupils and teacher but that there is no real evidence that this course, or one like it, should replace the standard tenth grade geometry course. On the other hand, there is limited evidence that some pupils benefit from a more algebraic course, and perhaps the Wesleyan experiment described by Mr. Sitomer will throw more light on the subject.

The experiment was funded by the Marcell M. Holzer Fund, and directed by Howard Fehr of Teachers College. Howard Levi wrote the materials, and the author of this article taught the course.

A course based on the same material was discussed in this section by Harry Sitomer in this MONTHLY, 72 (1965) 416–417.

Reference

1. Dieudonné's remarks in his "Euclid must go" speech in *New thinking in school mathematics*, OEEC, Paris, 1961.

A CASE FOR APPLICATIONS OF LINEAR ALGEBRA AND GROUP THEORY

R. K. JARVIS, Groton School, Massachusetts

This is an appeal for help. Briefly, I am a physicist with the usual sort of background in mathematics, but teaching mathematics extensively at the high-school level. I make no claims as a mathematician, nor many as a physicist, but I have had extensive experience in teaching mathematics and science to high-school students and college freshmen.

My problem concerns the applications of mathematics in general, but in particular the applications of linear algebra and group theory. To particularise still further, it is especially the application of matrices in linear algebra which concerns me most. It is a matter of the utmost concern to me, in whatever I teach, to show the power of mathematics in addition to its inherent fascination and beauty, and I am always very grateful to hear from any professional, whatever may be his field, of some new and interesting way in which he has put some mathematics to work. Many of the applications most commonly exhibited at the school level are certainly dated and new ones are badly needed.

In linear algebra, especially matrices, which has just begun to creep down into the schools in a significant way, I have been unable to find more than a very few significant applications which do not require extensive training in other disciplines in order to be appreciated. Perhaps I am asking too much and it is indeed impossible to satisfy the requirements of real significance and comprehension without extensive extra-mathematical training, but I hope not. There are a great many very bright men and women in science and applied mathematics, and matrices are extensively used. I have been encouraged to think by one university professor, who unfortunately does not have the time to work on this himself, that such applications as I could make use of in teaching bright high-school students are available. My question is, "Where?"

Two or three mathematicians I have spoken to think I am too concerned with the applications. "The subject stands on its own feet." It is not that I do not appreciate this point of view, but there are high-school and beginning college students, too intelligent to be ignored, for whom "pure mathematics is to applied mathematics, as crossword puzzles are to literature." A mathematician may disagree with this attitude, but it cannot be proved wrong. I recently attended a set of undergraduate lectures and it was clear both during class and from conversation afterwards, that a large percentage of the audience were left untouched, were baffled and frustrated by what for them was not meaningful or useful and purely academic. I have been approached by a few undergraduates and some of the boys I teach, all of them doing linear algebra, with the same

complaint; "So what?" I think that the fact has to be accepted that not all students are inspired by linear algebra *per se* and their lack of appreciation is not overcome by the all too numerous purely abstract presentations of the subject in current texts. If I may quote one professor, albeit talking of some presentations of the calculus: "Too much purity does not lead to fertility." There are many potential users of mathematics passing through teachers' hands, who need to see mathematics work if they are to apply it. I reject the idea that if any student cannot appreciate the linear algebra of itself then that is his own bad luck. Mathematicians who insist on the "pure Approach" and the pure approach alone for their recommendations as to what should be taught, particularly at my level, are, in my opinion, doing this country a great disservice which could have tragic consequences. There always will be more users than creators of mathematics and using mathematics does not simply come naturally once a student has some mathematics. He needs to be trained. To be sure he should understand what is mathematics and what is not in any application, but that is another problem. Men and women with a real command of mathematics, whether their viewpoint is pure or applied, are too valuable for us to reject any, simply because they are asking for a different point of view.

Perhaps people may ask "Why worry with linear algebra at the high-school level anyway?" My reason comes from a few very able students, now on their way to degrees and graduate degrees in science who insist that I should teach some. They claim that an early introduction to it would have been a great help to them. Who am I to argue? They are a part of the new generation of scientists, applied mathematicians and engineers and they ought to know.

You may ask what I regard as a significant application of mathematics. I draw the following from physics for obvious reasons but I have no prejudice as to field of application, scientific or otherwise. This example does not need extensive preparation in another field. A simple harmonic motion is not difficult to understand nor to demonstrate. Most twelfth grade students have dealt with Newton's law in their physics, or even if they have not, $F=ma$ is not difficult to accept on an intuitive basis, particularly at this age. It needs no detailed explanation to describe how the equation of a simple harmonic motion follows from here with a little integration, but that is not all. One of the best devices in the teaching of mathematics is prediction, and consideration of a point moving subject to two such motions superposed at right angles allows one to predict its paths given certain changes in amplitude, frequency and phase. Now show precisely this using an oscilloscope and, judging by reactions I have had, this is a really significant application. There is some excitement, there is some suspense, there is a sense of accomplishment.

And so I make an appeal to anyone who can find a moment, to do a little hard thinking and come up with some problems they have dealt with, particularly ones where matrices were used, which will show the power of the mathematics. The problems should be easy to introduce without extensive digression into other fields (it would be unrealistic to expect that no digression will be

necessary) and, if possible, should create some excitement. Matrices, one tells the student, are indispensable in many fields to avoid a quagmire of symbols. Can this be shown in any reasonably easily understood application? How can I convey to the beginning student in linear algebra that, beautiful as it is mathematically, the algebra of matrices is much more than an interesting mental exercise, indeed more than one of the foundation stones of mathematics? References to sources would be useful, but better still, something from a person's own work, maybe just a small piece of something much larger, but thought out in some detail as to introduction, significance and possible presentation. I hope that this is not so much of a chore that no one will feel inclined to undertake it. Will anyone help?

AN EXPERIMENT IN TEACHING MATHEMATICS AT THE COLLEGE LEVEL BY PROGRAMMED INSTRUCTION

T. A. DAVIS, DePauw University

The experiment reported here was concerned with comparing the effectiveness of programmed instruction with that of the traditional lecture and textbook approach in certain portions of our beginning course in analytic geometry and calculus.

The report is divided into five sections: (1) the reason we felt the need for programmed material, (2) the contents of the program, (3) the method of conducting the comparison, (4) the results, and (5) the conclusions drawn from the experiment.

(1) Many high schools now offer a course in analytic geometry or precalculus mathematics, but others do not. Hence the students taking analytic geometry and calculus at the college level have a wide range of previous training. If the traditional lecture and discussion technique is used to review this material, many of the best prepared students become bored and lose interest, while those with little or no preparation become lost and discouraged.

It was our hope that the students could learn this pre-calculus material more efficiently by using programmed instruction since the program would allow each student to spend as much time on each topic as he needed.

We also hoped that the use of the program would free valuable class time for topics in calculus.

(2) The program contained the material usually found in the first 50 to 60 pages of college textbooks on analytic geometry and calculus: coordinates, inequalities, absolute values, directed distance, distance in the plane, midpoint formula, slope of a line, parallel and perpendicular lines, the angle between two lines, graphs and equations, intercepts, symmetry and asymptotes, and the equations of the straight line.

The program was written by the author pursuant to a contract with the United States Office of Education, Department of Health, Education and Welfare.

(3) The experiment was conducted during the first semester of the 1964–65 academic year. Five sections of our beginning course in analytic geometry and calculus and three teachers were involved. The sections were not chosen at random. Each student signed up for whichever section he preferred. But there was no appreciable difference between the mean of the scholastic aptitude test scores of the programmed and nonprogrammed groups. (See Table I.) Three of the sections, 2, 4, and 5, each with a different instructor, had lectures and discussion as usual over the material covered by the program. The other two sections, 1 and 3, began immediately with functions and limits. At the same time, these students began working through the program on their own.

Since we planned to take three or four weeks of classroom time to cover the material in the program, we required the two sections using the program to complete it in four weeks. And, since the programmed sections were doing about one-fourth of the semester's course work on their own, they met only three days a week during the semester, while the other three sections met four days a week.

Each of the five sections was given a pretest on the programmed material on the first day of the semester. Four weeks later, each was given a posttest.

In addition to the common pretest and posttest, another hour test on topics in calculus and the final examination in the course were given in common to all five sections. The students were classified as A, B, C, D or F students on the basis of this common semester total.

(4) A summary of the results is given in Table I below.

TABLE I

Section	1 *Program*	2	3 *Program*	4	5
Number of students	25	26	27	29	21
SAT Verbal Mean	589.4	583.5	575.6	566.0	601.2
SAT Math Mean	667.3	684.8	680.9	657.9	695.6
Pretest Mean	31.2	34.9	36.5	31.6	38.4
Posttest Mean	88.3	87.3	91.7	86.0	84.3
Gainscore* Mean	57.1	52.4	55.2	54.4	45.9
Semester Average Mean	76.6	77.1	78.9	70.1	76.2

* Gainscore = Posttest − Pretest

TABLE II (Semester Scores vs. Posttest Scores)

POSTTEST SCORES	F Program	F non prog.	D Prog.	D non prog.	C Prog.	C non prog.	B Prog.	B non prog.	A Prog.	A non prog.
A (above 1 S.D.; to 100%)					3 (13.0%)		5 (55.6%)	1 (9.1%)	5 (50.0%)	4 (30.8%)
B (½ S.D. to 1 S.D.)	1 (16.7%)	1 (8.3%)	1 (25.0%)		7 (30.4%)	5 (17.9%)	3 (33.3%)	4 (36.4%)	4 (40.0%)	8 (61.5%)
C (−½ S.D. to ½ S.D.)	2 (33.3%)	5 (41.7%)		3 (25.0%)	11 (47.8%)	14 (50.0%)	1 (11.1%)	5 (45.4%)	1 (10.0%)	1 (7.7%)
D (−1 S.D. to −½ S.D.)		2 (16.7%)		5 (41.7%)	1 (4.4%)	3 (10.7%)		1 (9.1%)		
F (below −1 S.D.)	3 (50.0%)	4 (33.3%)	3 (75.0%)	4 (33.3%)	1 (4.4%)	6 (21.4%)				

0 F −1 S.D. D −½ S.D. C ½ S.D. B 1 S.D. A 100%

SEMESTER SCORES

(5) Conclusions:

(a) The program achieved the desired results. That is, the sections that used the program did as well as those who had the usual lecture and discussion approach. In fact, the program people did slightly better.

(b) The use of the program also enabled the students to save a considerable amount of time. We asked the students to keep a record of how long they took to work through the program. The time spent ranged from 8 to 20 hours with an average of 12 hours.

On the other hand, the non-programmed sections spent from 12 to 16 classroom hours on the same material. And, if it is assumed that the time spent on homework was up to two hours a day, a non-program student could have spent as much as 48 hours.

(c) The mean $\overline{X}_s$ for the semester averages was 75.94 with a standard deviation of 12.18. For convenience we divided the students into five categories. A students were those with scores above 1 S.D., B students had scores between

$\frac{1}{2}$ S.D. and 1 S.D., C students had scores between $\pm\frac{1}{2}$ S.D., D students had scores between -1 S.D. and $-\frac{1}{2}$ S.D., and F students were those with scores below -1 S.D.

Similarly the posttest scores were divided into 5 classes using the mean $\overline{X}_p=87.79$ and the standard deviation 8.12.

Table II gives the performance of the A, B, C, D and F students for both the programmed and non-programmed sections.

For example, of the 23 C-students who used the program, 3 of them or 13% did A work on the Posttest, 30.4% did B work on the posttest, 47.8% did C work, 4.4% did D work, and 4.4% failed.

On the other hand, of the 28 C students in the control groups, none did A work on the posttest, 5 or 17.9% did B work, 50% did C work, 10.7% did D work, and 21.4% failed.

We can make the following observations from Table II.

(i) The F and D students are unpredictable, but of the two groups, program and non-program, a higher percentage of the non-program students did better on the posttest than for the semester as a whole. Thus it would appear, on the basis of this small sample, that an F or D student has a better chance of doing well if he does *not* use the program.

(ii) On the other hand the C, B and A students who used the program did better as a group on the posttest than they did for the semester as a whole, while the control group did less well. Thus the A, B and C students seem to profit from using the program.

(d) While most of the program users did as well or better than their non-program counterparts, there were one or two of them who did not want to use programmed learning and who did badly when asked to use it.

(e) In summary, we find that the program teaches this preliminary material more efficiently than can be done by the traditional classroom approach.

A MATHEMATICS FILM SERIES

FRANCIS SCHEID, Boston University

The preparation of 120 mathematics films for the Harvard Commission on Extension Courses has now been completed. This work was sponsored by the Special Projects Division of the U. S. Navy and was done at the WGBH-TV studios in Boston. The films have been used primarily by the crews of Polaris submarines while on patrol [1], but are now being presented to selected units of the surface fleet. The films have also been broadcast over various educational TV channels. Of eight units, seven were prepared by the author and one (Statistics) by Prof. Gottfried Noether, also of the Boston University faculty. The policy in preparing each unit was to try to be futuristic. As a result the first three units are currently being presented to selected ninth grade students of the Latin schools in Boston, as an educational research project, to be reported later

under the title "Towards a Ninth Grade Calculus." Descriptions of the last six units follow. Descriptions of units 1 and 2 appear in [1].

3. Introduction to calculus. The principal objective of this series is to show how calculus grows directly out of arithmetic, both being parts of the same "game of numbers." The central idea is, therefore, *number*, and all other ideas presented are developed in terms of number. The series opens by introducing the idea of *function* as a set of X, Y number pairs. Starting with simple functions, involving only a few number pairs, it is shown how the physical idea of speed and the geometrical idea of steepness led earlier mathematicians to more sophisticated functions, and to the idea of *derivative*. Newton's original work on derivatives, and Berkeley's criticisms of its logical honesty, are compared with a modern treatment. This serves to emphasize the importance of logical honesty in mathematics, and also helps to clarify what is basically a sophisticated idea. The constant, linear, and quadratic functions then receive special attention. Their theory is developed and numerous applications are offered. The series presupposes a modest understanding of algebra and coordinate geometry, and since it focuses on the relationship between arithmetic and calculus, will be of interest to teachers at all levels as well as to prospective engineers or scientists.

4. The power functions. This series focuses on the idea of mathematical *power*, its theory and application. This is a famous and useful part of the "game of numbers." Here the algebra and calculus of the power functions are developed simultaneously. After the simplest integral powers (1, 2, 3, etc.), polynomials are studied and applied to problems of bending, smoothing, trajectories, etc. Reciprocal powers and rational functions (quotients of polynomials) come next. The crucial idea of inverse functions is then introduced. It is first illustrated by the square-root function, but then becomes the path to a more sophisticated power concept. A logarithm function is defined, and "powers of e" enter in the inverse of this logarithm function, followed by a general definition of "X to power P." The use of the word power in this broader sense is then justified by showing that the main theory of simple, integral powers remains true for the more sophisticated powers also. Applications to problems of population growth, pursuit, atmospheric drag, and numerous other fields are included. The series presupposes a brief introduction to calculus.

5. The trigonometric functions. This series offers an overall view of "trigonometry," from its *ancient origins* to the modern theory. It begins with Thales' alleged computation of the height of the Great Pyramid and Erathosthenes' computation of the radius of the earth. Problems of this sort led our mathematical ancestors to such ideas as angle, triangle, sine, cosine, tangent, arc of circle, arcsine and so on. Important problems of the same sort arise today and are solved in the same way. The emphasis in such applications is on the geometry. Sines and cosines are thought of as quotients of side-lengths of triangles. But sines and cosines are also numbers, and the question of how they could be given

strictly arithmetical definition and made honest, logical parts of the game of numbers was eventually raised. This proved to be a difficult question, but a *breakthrough* was finally achieved through a study of the rate at which angles and sines change relative to each other. But rates of change suggest calculus, and calculus definitely dominates the action from this point onwards. An "arcsine function" is defined as an integral. Sines then appear in the inverse function, and cosines and tangents easily follow. The *modern theory* of the trigonometric functions is then based on this strictly arithmetical approach. Though details of proofs are frequently omitted, the path to honest logic is mapped with care, so that the theoretical structure of this part of the game of numbers can be clearly seen. The fact that this theory is useful is made evident by the increasing complexity of the applications treated, chosen mainly from the fields of geometry and physics. Theory also leads to increasingly powerful methods for computing approximations of sines and cosines, as well as of the number pi which is so heavily involved. The theme of the series is "how trigonometry grew from ancient geometrical origins to an honest, logical part of the game of numbers, and how much its area of application expanded in the process."

6. Introduction to statistics. Statistical testing and estimation from a non-parametric or "distribution free" point of view are emphasized. The course includes elementary discussions on populations and samples, the idea of the "probability of an event," and the concept of independent and dependent events. The binomial and normal distributions are introduced both for their intrinsic value to classical statistics and because of their usefulness in non-parametric statistics. Point and interval estimation are considered for the binomial distribution, and this leads to the development of statistical tests and point an interval estimation procedures which are not dependent on the underlying probability distributions. Some such tests and estimation procedures considered are the sign test, the Wilcoxon tests, the χ^2-test for goodness of fit, and Kendall's T for association. The course assumes a background in high school algebra, and is aimed at those unacquainted with statistical tehniques as well as those who have had some classical statistics and are interested in seeing non-parametric alternatives to classical methods.

7. Introduction to modern algebra. The main objective of this course is to point out the nature of both abstract and applied mathematics. For this purpose two algebraic structures are compared. The first is Boolean algebra. It is presented axiomatically, several of its major theorems being proved. It is then applied in the three classical ways, to set theory, to the logic of statements, and to electrical machines. The second structure studied is the field. Its axioms and theory are compared with Boolean counterparts. Finite fields are applied to problems of experimental design. Finally binary symbols are used to connect the two structures, and a Boolean machine is designed capable of computing field sums. A typical program for such a machine is analyzed in detail and completes the unit.

8. Probability. This course begins with a careful treatment of combinatorial problems, emphasizing the enumeration of arrangements, combinations and partitions. The tree method is the principal tool, but recursions are also introduced whenever suitable. Probability is then presented as a mathematical abstraction, with axioms and theory. The tree method is again emphasized. Applications of various origins are presented, including biological, economic, gambling and military.

References

1. Preliminary report, "The Polaris Mathematics Program," this MONTHLY, 71 (1964) 422–424.
2. For units 1 to 4 mimeographed materials are available. Publication will follow.
3. For units 7 and 8 see "Elements of Finite Mathematics" by Scheid; Addison-Wesley. Reading, Mass., 1962.

BRIEF COMMENT

The Immutable Ph.D., EVERETT WALTERS, *Saturday Review*, January 15, 1966, No. 3, 49 (1966) 62 ff.

This article treats of some of the problems confronting the graduate schools and the Ph.D. degree. "Despite a century of criticism, and of dramatic change in other facets of education, little change has been made in the Ph.D. pattern since 1861 when Yale awarded the first academic doctorates in the United States. American graduate schools still hold to the traditional requirements for the degree in the face of new demands for doctors as research scientists, college teachers, business leaders, government officials, and continued criticism of the programs that lead to the degree." A degree intermediate between the Ph.D. and the MA is suggested, the notion being similar to the Doctor of Arts degree suggested in recent years by some members of the mathematical community. However, "it seems safe to say that the Ph.D. pattern, and such other doctoral patterns as education (Ed.D.), business (D.B.A.), and music (Mus.D.), will be held to with only minor genuflections to the problems of college teaching. And it appears that despite the greatly increased number of doctors turned out by the nation's graduate schools, there will not be a sufficient number for all institutions. Obviously the major universities and the affluent colleges will be able to attract the doctors, but the small four-year colleges and junior colleges will not be so fortunate. The latter must be content with masters and bachelors."

The Case of the Lingering Degree, FREDERIC W. NESS, *Saturday Review*, January 15, 1966, No. 3, 49 (1966) 64 ff.

This article discusses some of the problems, both academic and social, of the student who has completed all his work toward the Ph.D. except for the dissertation. This group of students seems to fall into three categories: the professional graduate student, the perpetual part-time graduate student, and the student who through force of economic necessity has had to interrupt his graduate study at this crucial point.

The author stresses the need for creative solutions to the problems of the ABD ("All But Dissertation") student and urges that graduate faculties attack these problems, for "the regrettable fact is that when a problem with serious social implications cannot be solved from within, it will be solved from without, often with some less desirable result. For example, there has just been introduced to the California Legislature a bill calling for a new type of doctorate to meet the state's need for teachers in its burgeoning system of higher education. Legislators will act if educators will not, where the public interest is involved."

Analysis of Research in The Teaching of Mathematics, KENNETH E. BROWN AND THEODORE L. ABELL, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 1965 (Catalog No. FS 5.229:29007-62) vii+99 pgs, 45 cents.

A review of research in the teaching of mathematics at all levels—elementary, junior and senior high school, and college—which was largely undertaken in the years 1961–1963. Besides giving detailed descriptive listings of the individual research projects, many of which are either doctoral or masters dissertations, the report organizes the individual studies into groups as they have bearing on specific questions of mathematical education. For example, such questions as "To what extent can mathematics concepts be developed in the elementary grades?," "Do manipulative materials increase achievement?", "Is programed instruction effective?", "Does study of calculus in high school affect achievement in college calculus?", "Is televised instruction effective?", and "How can success in college mathematics be predicted?" are mentioned and the studies bearing upon them cited. A list of as yet unanswered questions for possible future studies is given, and it is recommended that crucial problems in mathematical education be identified, that they be treated by teams of researchers rather than individuals, and the results be clearly reported and that the reports be given wide dissemination.

Advanced Symmetry Concepts in Grade 6, JEAN M. MILLIREN, *Programed Instruction*, 5 (1965) p. 1 ff.

This is a report of a project in teaching notions of geometric symmetry to a group of seven sixth-graders via programmed material. The main interest of the report can be gleaned from the following: "As we reviewed concepts, I took an active part in providing repeated opportunities for the students to express their new knowledge. Far from feeling supplanted by a programmed teaching sequence, I feel that my own role gained in importance. It was a most successful way to bring enrichment to an accelerated group." "Creativity did not suffer." "Within the group I sensed an augmented desire to discover new ideas, to experiment and to inquire."

All this suggests that with appropriate use of programmed material some of the fears associated with such material may never be realized. It is frequently mentioned that teachers do not like programmed material because they fear being replaced by a machine and that programmed material is dehumanizing—

that it leads to a stifling of creativity. It appears that with skillful use of programmed materials, exactly the opposite can occur.

The Teaching of Sets in Schools, K. R. McLean, London, 1965, viii+46 pgs.

This pamphlet was sponsored by the Modern Trends Sub-committee of the Teaching Committee of the Mathematical Association and was published by the Mathematical Association. The table of contents lists among other chapters, the following: Sets, Logic, Circuits, Boolean Algebra, Sets and Elementary Probability, The Idea of a Function, and The Use of Set Language in Algebra and Geometry. The material in these sections is detailed and indicates some of the things that may be done with sets and set theory at the school level. All the material has been tried at schools in Britain and various difficulties as well as possibilities have been included. The introduction contains a discussion of the advisability of introducing the notions of set theory at the school level and gives some reasons for so doing.

For the secondary school teacher in this country interested in introducing such material in his classes, the pamphlet should be valuable in giving him some ideas already used successfully elsewhere.

PROBLEMS AND SOLUTIONS

Edited by E. P. Starke

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to M. S. Klamkin, Ford Sci. Lab., P.O. Box 2053, Dearborn, MI 48121. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before October 31, 1966.

E 1895. *Proposed by A. A. Gioia, Texas Technological College*

Let a complex-valued arithmetic function f be called

(1) additive if $f(a+b)=f(a)+f(b)$,

(2) weakly multiplicative if $f(ab)=f(a)f(b)$ whenever $(a, b)=1$,

(3) strongly multiplicative if $f(ab)=f(a)f(b)$.

Prove that if a function is additive and weakly multiplicative, then it is strongly multiplicative.

E 1896. *Proposed by David Singmaster, University of California, Berkeley*

For every prime p show that

$$\binom{np}{p} \equiv n \pmod{p^2}.$$

E 1897. *Proposed by F. P. Callahan, Blue Bell, Penna.*

A bug is crawling on the edges of a regular dodecahedron. Each time it comes to a vertex it chooses with equal probability one of the two edges which end in that vertex distinct from the one by which it reached the vertex. What is the expected distance it covers in order to get from a vertex A to any other vertex B? ($A = B$ is not excluded.) On starting, the bug can go in any of the three directions. (Cf. E 1752 [1966, 200].)

E 1898. *Proposed by T. L. Saaty, U. S. Arms Control and Disarmament Agency*

Give an elementary proof that the volume of the region in E_n enclosed by $\sum_{i=1}^{k} x_i \geqq 0$, $k=1, \cdots, n$, $|x_i| \leqq \frac{1}{2}$, $i=1, \cdots, n$ is given by $(-1)^n\binom{-1/2}{n}$.

E 1899. *Proposed by C. R. Knapp, University of Kansas*

Let $a_j = e^{2\pi j i/n}$, $i = (-1)^{1/2}$. Further, let a_k be any primitive nth root of unity. Prove that

$$a_k^{k^{\phi(n)-1}} = a_1, \qquad j = 1, 2, \cdots, n,$$

where $\phi(n)$ is the Euler totient function.

E 1900. *Proposed by D. R. Rao, Secunderabad, India*

Find the solution of the system $X + Y = Z + U = W^2$, $X^2 + Y^2 = Z^2 - U^2$ where X, Y, Z, U, W are positive integers.

E 1901. *Proposed by Richard Laatsch, Miami University, Oxford, Ohio*

Let $\triangle ABC$ have unit area. Construct $\triangle A'B'C'$ such that each of A, B, C is an interior point on a different one of its sides. (Label so that A and A' are on opposite sides of BC, etc.) Find the supremum and infimum of the product of the areas of $\triangle A'BC$, $\triangle AB'C$, and $\triangle ABC'$ if:

1. Two sides of $\triangle ABC$ are parallel to sides of $\triangle A'B'C'$;
2. One side of $\triangle ABC$ is parallel to a side of $\triangle A'B'C'$;
3. No parallelism restriction is made.

E 1902. *Proposed by David Shelupsky, the City College, New York*

Let $s(x) = \sum_{m=0}^{\infty} x^{2m+1}/1\cdot 3 \cdots (2m+1)$. Prove that for all nonnegative x

the inequality

$$e^{-x^2/2}s(x) < (\tfrac{1}{2}\pi)^{1/2}$$

holds, and that as $x \to +\infty$ the left-hand side is monotone increasing with the right-hand side as its limit.

E 1903. *Proposed by George Eldredge, El Cerrito, California*

Let an n-ladder of twos, L_n, be defined as follows:

$$L_n = 2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$$

where there are n twos. Let N_n be the number of distinct integers that can be obtained from L_n by the appropriate insertion of a set of unambiguous nested parentheses. For example, $N_3 = 1$; $N_4 = 2$. Find N_n.

E 1904. *Proposed by D. L. Silverman, Beverly Hills, California*

Solve in positive integers: $x+y+z = xyz(x+2y-2z-2)$.

SOLUTIONS OF ELEMENTARY PROBLEMS

On the Representation of Integers

E 1776 [1965, 420]. *Proposed by Carolyn C. Styles, San Diego Mesa College*

For any even integer m, $m>2$, there is an infinity of odd integers not the sum of a prime and a positive power (>1) of m.

I. *Solution by D. I. A. Cohen, Princeton University.* With $k>0$, the odd number $m(m-1)^k+1$ may not be written as $p+m^b$, p a prime, $b \geqq 1$. Recalling that $m>2$, it follows from $m(m-1)^k+1 = p+m^b$ that $p \equiv 1 \pmod{m}$ and that $(m-1)$ divides p, i.e., $p = m-1$, an impossible situation.

II. *Solution by R. B. Eggleton, Avondale College, Cooranbong, N.S.W., Australia.* More generally, we show that for any integer $m \geqq 3$, there is an infinity of integers of parity opposite to that of m, which cannot be represented as the sum of a prime and a positive power of m.

Proof. Let a be the smallest odd divisor (>1) of m if m is odd, or of $m-1$ if m is even. With $r \geqq 2$, let $b(s, k) = m^r - m^s + (2k+1)a$, where $k(\geqq 1)$ satisfies $(2k+1)a < m^r(m-1)$ and $s(\geqq 1)$ satisfies $b(s, k) > 0$. This implies $1 \leqq s \leqq r$. Also $a \mid b(s, k)$ since $m(m-1) \mid (m^r - m^s)$, and since $b(s, k) > a$ it follows that a is a proper factor of $b(s, k)$. That is, $b(s, k)$ is composite; clearly it is also odd. Thus, $c(r, k) = m^r + (2k+1)a$ can only be expressed in the form $m^s + b$ with composite b, and there is an infinity of such numbers c.

Also solved by J. L. Brown, Jr., E. O. Buchman, D. M. Hancasky, D.C. B. Marsh, Robin Sibson (England), and Guy Torchinelli.

$$\{(a, c) \mid a, c > 0, a^n n! > cn^n \forall n\}$$

E 1777 [1965, 420]. *Proposed by Horváth Sándor, University of Technology, Budapest, Hungary*

For which positive values of a and c is $a^n n! > cn^n$ true for every positive integer n?

Solution by Walter O. Egerland, Edgewood Arsenal, Md. The proposed inequality is true for every positive integer n for all positive a and c which satisfy $a \geqq e$ and $c < a$. To see this, put $s_n = a^n n^{-n} n!$ $(n = 1, 2, \cdots)$; then $s_n = a^{-1}(1+1/n)^n s_{n+1}$. If $a \geqq e$, s_n increases monotonically so that $s_n > c$ holds for every positive integer if and only if $s_1 = a > c$. If $0 < a < e$, s_n decreases monotonically for $n \geqq N(a)$ and converges to zero by Stirling's formula so that $s_n > c > 0$ cannot hold for any n.

Also solved by A. N. Aheart, J. W. Baldwin, Y. M. ben-David, P. J. Campbell, D. I. A. Cohen, David Cohoon, R. B. Eggleton (Australia), J. A. Faucher, Robert Feinerman, N. J. Fine, Michael Goldberg, P. L. Gore, D. M. Hancasky, Erwin Just, E. S. Langford, D. C. B. Marsh, D. E. Nixon, C. B. A. Peck, J. M. Perry, P. A. Scheinok, Ralph Schreiber, Robin Sibson (England), Al Somayajulu, Sidney Spital, Gomer Thomas, Rory Thompson, W. C. Waterhouse, B. D. Wick, J. M. Wild, Jr., and K. L. Yocom.

The Radii of a Triangle

E 1778 [1965, 420]. *Proposed by Hüseyin Demir, Middle East Technical University, Ankara, Turkey*

If R, r, r_1, r_2, r_3 are the circumradius, inradius and exradii of a triangle, prove that

$$\frac{1}{r^3} - \frac{1}{r_1^3} - \frac{1}{r_2^3} - \frac{1}{r_3^3} = \frac{12R}{r \cdot r_1 \cdot r_2 \cdot r_3}.$$

Solution by Ralph Schreiber, Warsaw High School, Warsaw, Ind. Denote by $\triangle$ the area of triangle ABC, by s the semiperimeter, by r_a the exradius corresponding to side a, and so forth. We recall familiar identities:

$$\triangle = rs = r_a(s-a) = r_b(s-b) = r_c(s-c) = \sqrt{rr_ar_br_c} = abc/4R.$$

Thus

$$\frac{1}{r^3} - \frac{1}{r_a^3} - \frac{1}{r_b^3} - \frac{1}{r_c^3} = \frac{1}{\triangle^3}\left[s^3 - (s-a)^3 - (s-b)^3 - (s-c)^3\right]$$
$$= 3abc/\triangle^3 = 12R/\triangle^2 = 12R/rr_ar_br_c.$$

Also solved by A. N. Aheart, Leon Bankoff, W. J. Blundon, D. I. A. Cohen, Ragnar Dybvik (Norway), Mrs. A. C. Garstang, Michael Goldberg, Louise S. Grinstein, D. M. Hancasky, E. S. Langford, Ruth S. Lefkowitz, F. Leuenberger (Switzerland), Andrzej Makowski (Poland), D. C. B. Marsh, F. R. Prieto, J. M. Quoniam (France), S. Bhaskara Rao (India), Simeon Reich (Israel), P. A. Scheinok, Klaus Schmitt, R. Sivaramakrishnan (India), Sidney Spital, Sister M. Stephanie, M. V. Tamhankar & M. B. Suryanarayana (India), Simon Vatriquant (Belgium), C. S. Venkataraman (India), William Wernick, and the proposer.

The Altitudes and Exradii of a Triangle

E 1779 [1965, 420]. *Proposed by Hüseyin Demir, Middle East Technical University, Ankara, Turkey*

If h_i and r_i are the altitudes and exradii of a triangle prove that

$$\frac{r_1}{h_1}+\frac{r_2}{h_2}+\frac{r_3}{h_3} \geqq 3.$$

I. *Solution by D. C. B. Marsh, Colorado School of Mines.* Since $2/h_i = 1/r_j + 1/r_k$ for i, j, k a cyclic permutation of 1, 2, 3 (R. A. Johnson, *Modern Geometry*, p. 189; or use the identities in E 1778 above together with $1/r = 1/r_1 + 1/r_2 + 1/r_3$) it follows immediately that $\sum(r_i/h_i) = \frac{1}{2}\sum_{i\neq j}(r_i/r_j) \geqq \frac{1}{2}(6) = 3$, since $(x/y)+(y/x) \geqq 2$ for positive x, y. Moreover, equality obtains only if $r_1 = r_2 = r_3$, i.e., the triangle is equilateral.

II. *Solution by H. Guggenheimer, University of Minnesota.* We may generalize by securing the inequality $\sum(r_i/h_i)^n \geqq 3$, $n \geqq 1$. Actually more is true: Let t_i be the lengths of the angle bisectors of the triangle. Since $t_i \geqq h_i$, the proposed inequality is weaker than

$$\sum_{i=1}^{3}\left(\frac{r_i}{t_i}\right)^m \geqq 3 \qquad m > 0$$

which we now prove.

Let s be the semiperimeter of the triangle, a_i the sides. Leuenberger has proved (Elemente Math., 17 (1962) 45–46; see also 16 (1961) p. 129) that $t_i \leqq [s(s-a_i)]^{1/2}$. Hence

$$\sum\left(\frac{r_i}{t_i}\right)^m \geqq \sum_{i\neq j\neq k}\left[\frac{s(s-a_j)(s-a_k)}{s(s-a_i)^2}\right]^{m/2}$$

$$= \frac{1}{[(s-a_1)(s-a_2)(s-a_3)]^m}\sum_{j\neq k}[(s-a_j)(s-a_k)]^{3m/2}.$$

The desired result now follows from the geometric-arithmetic mean inequality, and again equality holds only for the equilateral triangle.

Also solved by A. N. Aheart, Leon Bankoff, W. J. Blundon, D. I. A. Cohen, Mrs. A. C. Garstang, Michael Goldberg, H. Guggenheimer, D. M. Hancasky, E. S. Langford, F. Leuenberger (Switzerland), Andrzej Makowski (Poland), F. R. Prieto, J. M. Quoniam (France), S. Bhaskara Rao (India), Simeon Reich (Israel), P. A. Scheinok, Ralph Schreiber, R. Sivaramakrishnan (India), Sidney Spital, Sister M. Stephanie, M. V. Tamhankar & M. B. Suryanarayana (India), P. D. Thomas, Simon Vatriquant (Belgium), C. S. Venkataraman (India), and the proposer.

Makowski's student, Tadeusz Figiel observed that the required inequality is equivalent to the fact that the area of an orthic triangle is not greater than one-quarter of the area of a given (acute-angled) triangle. [*Proof:* Let ABC be an orthic triangle of $A_1B_1C_1$. Then A_1, B_1, C_1 are the centers of ex-circles and the ratio of areas of ABC_1 and ABC is equal to the ratio of altitudes on the common side AB, i.e., r_3/h_3.]

Solution of a Trigonometric Equation

E 1780 [1965, 420]. *Proposed by Norman Brenner, Harvard University*

Find all pairs (α, β) of real numbers which satisfy the equation

$$\frac{\sin 2\alpha}{\sin(2\alpha + \beta)} = \frac{\sin 2\beta}{\sin(2\beta + \alpha)} .$$

Solution by D. C. B. Marsh, Colorado School of Mines. Let $\alpha+\beta=2s$ and $\alpha-\beta=2d$, transforming the given equation into

$$\frac{\sin(2s + 2d)}{\sin(3s + d)} = \frac{\sin(2s - 2d)}{\sin(3s - d)} . \tag{1}$$

Clear of fractions in (1), expand by sum and difference formulas and simplify, to obtain

$$\sin 2s \cos 3s \sin d \cos 2d = \sin 3s \cos 2s \sin 2d \cos d. \tag{2}$$

Product-to-sum formulas, with simplification, now produce

$$\sin s \sin 3d + \sin 5a \sin d = 0. \tag{3}$$

Finally, multiple-angle formulas reduce (3) to

$$\sin s \sin d(\cos^2 d + \cos s \cos 3s) = 0. \tag{4}$$

Examination of the factors in (4) enables us to list all real pairs (α, β) which satisfy the original equation:

I. From $\sin d=0$: $(\phi+2n\pi, \phi)$ where n is any integer and ϕ any real number except integral multiples of $\pi/3$ (to avoid 0 denominators);

II. From $\sin s=0$: $(-\phi+2n\pi, \phi)$, n integral and ϕ any real number except integral multiples of π;

III. From $\cos^2 d+\cos s \cos 3s=0$: $(s+d, s-d)$, where s acts as parameter and may be chosen as any real number in the intervals $(\pi/6+2k\pi, 5\pi/6+2k\pi)$, $(7\pi/6+2k\pi, 11\pi/6+2k\pi)$ with k integral, excepting s an odd integral multiple of $\pi/6$, $\pi/4$ or $\pi/2$; $d=\arccos(\pm\sqrt{(-\cos s \cos 3s)})$.

Also solved by E. S. Langford, and Simon Vatriquant (Belgium).

Continuous Functions with a Common Value

E 1781 [1965, 421]. *Proposed by Ray Redheffer, University of California, Los Angeles*

On an interval (a, b) let f and g be continuous, and f or g monotone. Suppose that some sequence $\{x_n\}$ satisfying $f(x_n)=g(x_{n-1})$ has a limit point on (a, b). Prove that then $f(x)=g(x)$ has a solution on (a, b).

I. *Solution by Sidney Spital, California State Polytechnic College.* Either $\{x_n\}$ is monotone, or it is not. If it is not, let f be monotone. Then $f(x_n)-g(x_n)$ $=f(x_n)-f(x_{n+1})$ must undergo sign changes between some members of $\{x_n\}$.

The requested result then follows from the continuity of f and g.

If $\{x_n\}$ is monotone its limit point on (a, b) becomes its limit. The desired result is now shown by taking the limit of $f(x_n)=g(x_{n-1})$ and again using the continuity of the two functions.

II. *Solution by T. J. Burke, New Jersey.* The conclusion follows without assuming the monotonicity of f or g. Suppose that $f(x)\neq g(x)$ for all $x\in(a, b)$. Then (without loss of generality) assume that $f>g$ on (a, b). (The alternative that $g>f$ on (a, b) can be handled by considering the functions $-f$ and $-g$.) Let x_0 be a limit point of $\{x_n\}$, $x_0\in(a, b)$. Then there is a subsequence $\{x_{n_k}\}$ such that $x_{n_k}\to x_0$. From $g(x_0)<f(x_0)$ it follows that there is an integer p such that, for any $k_1, k_2\geqq p$

$$(*)\qquad g(x_{n_{k_1}}) < f(x_{n_{k_2}}).$$

Now for some positive integer m, $n_{p+1}=n_p+m$; but

$$\begin{aligned} f(x_{n_{p+1}}) &= g(x_{n_{p+1}-1}) < f(x_{n_{p+1}-1}) \\ &= g(x_{n_{p+1}-2}) < f(x_{n_{p+1}-2}) < \cdots < f(x_{n_{p+1}-(m-1)}) = g(x_{n_p}), \end{aligned}$$

which contradicts (*). Hence the assumption $f>g$ on (a, b) must be false and the proof is complete.

Also solved by J. M. ben-David, E. O. Buchman, I. A. Cohen, N. J. Fine, D. A. Hejhal, James Humphreys & Ben Klein, James Joseph, Erwin Just, E. S. Langford, D. C. B. Marsh, J. D. Morrison, J. M. O'Neil, Paul Sontag, R. A. Struble, Gomer Thomas, and Maynard Tomer.

Several solvers assumed the convergence of $\{x_n\}$ without giving justification.

Coloring a Chessboard

E 1782 [1965, 421]. *Proposed by V. V. Menon, Indian Statistical Institute, Calcutta*

Given an $n\times n$ chessboard, a coloring of the board with k colors is defined as an assignment of k colors to the cells of the board such that (a) each cell is assigned a color, and (b) no two cells which are assigned the same color lie in the same row, same column, or same diagonal. What is the minimum number of colors required for a coloring?

Partial Solution by Michael Goldberg, Washington, D. C. Let k be the minimum number of required colors. Since there are n colors in any row, then $k\geqq n$. If $k=n$, then the n cells of any single color constitute a solution of the problem of n queens on an $n\times n$ board. See Ball and Coxeter, *Mathematical Recreations and Essays*, pp. 165–171; also Kraitchik, *Mathematical Recreations*, pp. 247–255. If the location of each queen can be obtained by a motion of $(1, v)$ from another queen (where v is prime to n), the solution is called regular. If the solution is regular (Kraitchik, p. 251), then the other colored sets can be obtained by changing the origin of the lattice. Regular solutions are always obtainable if and only if n is not a multiple of 2 or 3. Hence, if $n\equiv 1, 5 \pmod 6$, then $k=n$.

If $n \equiv 0$, 4 (mod 6), then $k = n+1$, since from each regular solution of an $(n+1)$ square in $n+1$ colors, one can obtain an n square in $n+1$ colors by removing a row and column associated with a corner cell.

If $n \equiv 3$ (mod 6), then $n < k \leqq n+2$, since a square of side $6a+3$ can be obtained by removing two rows and two columns from a square of side $6a+5$.

Similarly, if $n \equiv 2$ (mod 6), then $n < k \leqq n+3$, since a square of side $6a+2$ can be obtained by removing three rows and three columns.

Undoubtedly, individual and general results will improve these latter values. For example, if $n = 8$, then only 9 colors *ABCDEFGHJ* are needed as shown by the solution in which colors in the 8 rows are *ADHCFGBE*, *BEFGHCAD*, *CGBAEDHF*, *DHEFCAGB*, *ECDJGBFA*, *FAGBDHEC*, *GBCEAFDH*, *HFADBECJ*.

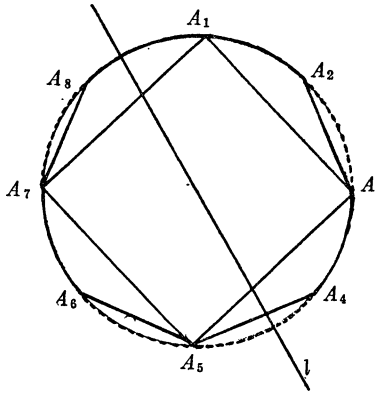
Also solved (partially) by Paul Sontag.

Area Bisectors of a Plane Convex Figure

E 1783 [1965, 421]. *Proposed by V. V. Menon, Indian Statistical Institute, Calcutta*

A figure is a compact convex set with nonempty interior in the Euclidean plane. A bisector of a figure is defined to be a straight line which divides the figure into two portions of equal area. Given $n (n \geqq 3)$, does there exist a figure such that through some point in it there pass exactly n (and no more) bisectors of the figure?

I. *Solution by D. I. A. Cohen, Princeton University.* Let $A_1, A_2, \cdots, A_{4n}$ be $4n$ equally spaced points on a circle. Draw line segments $A_{2m}A_{2m+1}$ for $m = 1, 2, \cdots, n$ and $A_{2m-1}A_{2m}$ for $m = n+1, \cdots, 2n$. Draw circular arcs $A_{2m-1}A_{2m}$ for $m = 1, \cdots, n$ and $A_{2m}A_{2m+1}$ for $m = n+1, \cdots, 2n$ $(A_{4n+1} = A_1)$. Consider the figure C whose boundary consists of the $2n$ segments and $2n$ arcs just described.



C is then the $2n$-gon $A_1A_3A_5 \cdots A_{4n-1}$ surmounted by $2n$ congruent "caps." (C will have the diagonal A_1A_{2n+1} as its only axis of symmetry.) Now we can prove that through the center of the circle the only lines that bisect the area of

C are the n diagonals of the $2n$-gon. Any other line through the center will have $n-1$ caps and half of the $2n$-gon on each side but will cut two caps unequally. (See figure drawn for $n=2$. Note that line l is not a bisector.)

II. *Solution by Michael Goldberg, Washington, D. C.* Yes. In my note *On Area-Bisectors of Plane Convex Sets*, this MONTHLY, 70 (1963) 529–531, it is shown that every regular polygon of $2k+1$ sides, or any closed convex non-circular curve which has $2k+1$ axes of symmetry which pass through the center, has the desired property. There is a central region in which every point has exactly $2k+1$ bisectors through it. Every point on the boundary of this central region (except the cusps) has exactly $2k$ bisectors through it.

Also solved by E. O. Buchman, H. Guggenheimer, and Andrzej Makowski (Poland).

Makowski cites an example given by E. Piegat (Wiad. Mat., 7 (1963) 51–56) which is essentially the same as I above.

Multiple Solutions of a Differential Equation

E 1784 [1965, 421]. *Proposed by A. W. Hales, Institute for Defense Analyses, and G. R. Sells, University of Minnesota*

Consider the first order differential equation $dy/dx=f(y)$, where f is continuous in a neighborhood of y_0. Assume that the solutions through y_0 are not unique, that is, assume that there are two solutions ϕ_1 and ϕ_2 such that $\phi_1(0)=\phi_2(0)=y_0$ and that ϕ_1 and ϕ_2 differ in every neighborhood of 0. Prove that $f(y_0)=0$.

Solution by N. J. Fine, the Pennsylvania State University. Suppose $f(y_0)\neq 0$. There is a neighborhood V of y_0 on which $f(y)$ is bounded away from 0. Let W be a neighborhood of 0 such that $\phi_1[W]\subset V$ and $\phi_2[W]\subset V$. For $x\in W$, define

$$H(x) = \int_{\phi_1(x)}^{\phi_2(x)} \frac{dy}{f(y)} \cdot \tag{1}$$

If $\phi_1(x)\neq\phi_2(x)$ for some $x\in W$, then $H(x)\neq 0$. But for all $x\in W$,

$$H'(x) = \frac{\phi_2'(x)}{f(\phi_2(x))} - \frac{\phi_1'(x)}{f(\phi_1(x))} = 0$$

by (1) and the differential equation. Therefore H is constant on W. Since $H(0)=0$, $H(x)=0$, $(x\in W)$. This contradiction completes the proof.

Also solved by Y. M. ben-David, E. O. Buchman, W. G. Dotson, Jr., D. A. Hejhal, M. E. Muldoon, Sidney Spital, R. A. Struble, D. M. Yasnyi, and the proposers.

Number of Elements in a Set of m-tuples

E 1785 [1965, 543]. *Proposed by Robert Bowen, University of California, Berkeley*

Let $T_m=\{(a_1,\cdots,a_m): a_i \text{ is a positive integer, } a_1=1,\ a_i\leqq 1+\max_{k<i} a_k\}$,

and $t(m)$ be the number of elements in T_m. Set $g_0(x)=1$ and $g_{k+1}(x)=(x-1)g_k(x)+g_k(x+1)$. Show that $t(m)=g_m(1)$. [Note correction in original statement: $g_1(x)=1$ is replaced by $g_0(x)=1$.]

Solution by E. O. Buchman, University of California, Los Angeles. We prove the

THEOREM. *The set* $S_{m,n}=\{(a_1, \cdots, a_m): a_i$ *positive integers,* $a_1\leqq n$, $a_i \leqq \max(1+\max_{k<i} a_k, n)\}$ *has* $g_m(n)$ *elements.*

Proof by induction on m. When $m=0$, the theorem is true if it is assumed that there is a unique null-sequence of length 0. When $m=1$, there are n sequences of length 1 in $S_{1,n}$; and $g_1(x)=x$, so $g_1(n)=n$.

For the case $m=j+1$, we assume the theorem for $m=j$. Then

$$S_{j+1,n} = \bigcup_{i=1}^{n} \{(a_1, \cdots, a_{j+1}) \in S_{j+1,n}: a_1 = i\} = \bigcup_{i=1}^{n} S_{j+1,n}^{(i)}.$$

If $i<n$, the possible sequences for $(a_2, \cdots, a_{j+1})$ come from the set $S_{j,n}$, so there are $g_j(n)$ elements in $S_{j+1,n}^{(i)}$. If $i=n$, the possible sequences for $(a_2, \cdots, a_{j+1})$ come from the set $S_{j,n+1}$, so there are $g_j(n+1)$ elements in $S_{j+1,n}^{(n)}$. Therefore $S_{j+1,n}$ contains $g_j(n)+g_j(n)+ \cdots +g_j(n)+g_j(n+1)$ elements, or $(n-1)g_j(n)+g_j(n+1)$ elements. By definition of g, $S_{j+1,n}$ contains $g_{j+1}(n)$ elements. The theorem follows by induction. The proposed problem specializes this result to $n=1$.

Also solved by D. C. B. Marsh, and by Stanton Philipp.

A Pavement of Tetrominoes

E 1786 [1965, 543]. *Proposed by Robert Spira, Duke University*

A rectangle can be made up of T-shaped tetrominoes (made of unit squares) if and only if each side of the rectangle is divisible by 4. (The simplest case is the 4×4.)

Editorial Note: Since the announcement of this problem, an independent proof by D. W. Walkup has been printed. See this MONTHLY [1965, 986–8].

Also solved by D. M. Hancasky.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz. Rutgers-The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before December 31, 1966.

5400. *Proposed by W. C. Waterhouse, Harvard University*

Find, or disprove the existence of, two (totally) ordered fields which are isomorphic as fields and isomorphic as ordered sets, but not isomorphic as ordered fields.

5401. *Proposed by Charles Wells, Duke University*

Let ϵ be a primitive nth root of unity. Define a matrix $A=(a_{rs})$ by $a_{rs}=\epsilon^r-\epsilon^s$ for $r, s=1, 2, \cdots, n$. Find the rank of A.

5402. *Proposed by David Shelupsky, The City College, New York*

For $N>-1$ let the functions f and g be defined by the expansions

$$f(x) = \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+N+1}}{(2m+N+1)!(2m+1)!},$$

$$g(x) = \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m+N}}{(2m+N)!(2m)!}$$

which somewhat resemble the expansions of the trigonometric functions. Prove that each of these functions has, in common with the trigonometric functions, an infinite number of distinct real zeros.

5403. *Proposed by Kwangil Koh, North Carolina State University*

It is known that a simple ring is not necessarily primitive. Prove that if R is a simple ring such that for each right (or left) ideal I of R, $IR=I(RI=I)$, then R is primitive.

5404. *Proposed by George Grossman, Board of Education, New York City*

By an almost-Pythagorean triangle we mean one whose sides are natural numbers satisfying $a^2+b^2=c^2\pm 1$. Derive formulas for determining a, b, c. In particular, show that if (a_1, b_1, c_1) is a solution, then there are infinitely many solutions (a, b, c) with $a/c=a_1/c_1$, and infinitely many with $a/b=a_1/b_1$.

5405. *Proposed by A. A. Mullin, University of California, Livermore*

Let ζ be Riemann's zeta function. Define $\zeta^2=\zeta(\zeta(\ ))$ and higher "powers" recursively. Prove that the set of all complex s for which $\lim_{n\to\infty}\zeta^n(s)$ exists is precisely countably infinite. Does there exist a nonreal s for which $\lim_{n\to\infty}\zeta^n(s)$ exists? What is the greatest real value of s for which $\lim_{n\to\infty}\zeta^n(s)$ exists?

5406. *Proposed by C. C. Lindner, Coker College, Hartsville, S. C.*

Prove that a finite group of order $2n$ which contains an element of order n has at least $\tau(n)+1$ normal subgroups. $\tau(n)$ is the number of divisors of n.

5407. *Proposed by S. W. Golomb, University of Southern California*

Let $f(n)$ be a function with both range and domain the positive integers, and define $g(n)$ to be the number of values of m such that $f(m)=n$. If $f(n)$ is monotone nondecreasing and satisfies $f(n)=g(n)$ for all n, then $f(n)$ is uniquely determined. Find an asymptotic expression for $f(n)$ as $n\to\infty$.

5408. *Proposed by Thaddeus Dankel, Princeton University*

Let C be any infinite collection of distinct Lebesgue-measurable subsets of

$[0, 1]$ each having measure $\geqq \epsilon$, where ϵ is a fixed positive number <1. Show that there exists an infinite subcollection $C_0 \subset C$ of distinct sets E such that $m(\bigcap_{E\in C_0} E) > 0$.

5409. *Proposed by R. A. Melter, University of Massachusetts*

Let k be a square-free positive integer greater than 1 and let R be a subring of the complete direct sum of an arbitrary number of copies of I_k, the ring of integers modulo k. Show that R is the direct sum of a finite number of p-rings. (A p-ring is a ring in which the identities $px=0$ and $x^p=x$ are satisfied for all x and for some positive prime p.)

SOLUTIONS OF ADVANCED PROBLEMS

Coregular Matrix Example

4538 [1953, 336]. *Proposed by Albert Wilansky, Lehigh University*

Given $\sum |b_k| < \infty$, must there exist a constant M such that whenever $\{x_n\}$, $n=1, 2, \cdots$, is a convergent sequence satisfying

$$\left| \sum_{k=1}^{n-1} b_k x_k + x_{n-1} + x_n \right| < 1$$

for all n, then $|\lim x_n| < M$?

Remark by the proposer. No solution to this problem was ever submitted to the MONTHLY. Because of this, the problem was published in the Research Problems section of the Bull. Amer. Math. Soc., 67 (1961) p. 355. Three solutions are now known to have been given. One, unpublished, by Lawrence Shepp, one by W. K. Hayman and A. Wilansky, Bull. Amer. Math. Soc., 67 (1961), pp. 554–555, and a generalization, by C-S Hsü, Proc. Amer. Math. Soc., 13 (1962) pp. 979–981.

On Finitely Generated Torsion Free Modules

5273 [1965, 323]. *Proposed by T. J. Head, Iowa State University*

Show that for a commutative integral domain R to have the property that each of its finitely generated torsion free modules is free it is necessary and sufficient that each finitely generated ideal of R be principal.

Solution by Gomer Thomas, University of Illinois.

Necessity. Let J be any finitely generated nonzero ideal of R. Since R has no zero divisors, J is a torsion free R-module and hence is (module) isomorphic to a direct sum of copies of R. Each of the summands of J is an ideal in R. No two nonzero ideals in R can intersect trivially, since then R would have zero divisors. Hence there is only one summand in J; i.e., $J \cong R$ (as R-modules). It follows that J is principal.

Sufficiency. Let M be a torsion free R-module generated by $m_1, m_2, \cdots, m_n$,

and proceed by induction on n. If $n=1$, then $M=Rm_1\cong R$. If $n>1$, let S be the pure submodule of M generated by m_n; i.e., $S=\{m\in M: rm=tm_n$ for some $r, t\in R\}$. Then M/S is torsion free with $n-1$ generators, so by the induction hypothesis M/S is free. M/S is then projective, so $M\cong S\oplus M/S$, and it is now sufficient to show that S is (module) isomorphic to R. As a direct summand of a finitely generated module, S is finitely generated, and the generators can be written (symbolically) as $(t_1/r_1)m_n, (t_2/r_2)m_n, \cdots, (t_n/r_n)m_n$. Then $r_1r_2\cdots r_nS$ is generated by $t_1'm_n, t_2'm_n, \cdots, t_n'm_n$ or $r_1r_2\cdots r_nS=Jm_n$, where J is the (principal) ideal in R generated by $t_1', t_2', \cdots, t_n'$. The result follows, since $S\cong r_1r_2\cdots r_nS\cong Jm_n\cong Rm_n\cong R$.

Also solved by V. C. Cateforis, E. R. Gentile (Argentina), Michael Rosen, P. J. Ryan, K. G. Wolfson, and the proposer.

Editorial Note. Cateforis and Gentile use theorem 2.4, p. 131 of Cartan-Eilenberg, *Homological Algebra*, to obtain an immediate proof of sufficiency, for which Wolfson cites the proof of the theorem that any finitely generated module over a commutative principal ideal domain is the direct sum of cyclic modules (p. 44 of I. Kaplansky, *Infinite Abelian Groups*).

Nilpotent Divisors of Zero

5277 [1965, 323]. *Proposed by Kwangil Koh, University of North Carolina*

It is known that in a Noetherian ring every irreducible ideal is primary (see, e.g., N. H. McCoy, *Rings and Ideals*, Carus Monograph No. 8, p. 200). Let R be a ring (not necessarily commutative) which satisfies the ascending chain condition on right ideals. Prove that if S is a two-sided ideal in R such that it is irreducible, in the sense that if J_i, $i=1, 2$, is a pair of right ideals such that $J_i\supset S$ then $J_1\cap J_2\supset S$, then each left divisor of zero in the ring R/S is nilpotent (where $\supset$ means "contains strictly").

Solution by Roger A. Abelsgaard, Bemidji State College, Minnesota. Suppose there is a divisor of zero in R/S, say $\alpha\in R/S$, $\beta\in R/S$ and $\beta\alpha=\bar{0}=S$ with $\alpha\neq S$ and $\beta^j\neq S$ for $j=1, 2, \cdots$. Then if we represent α as $a+S$ and β as $b+S$, we have $a\notin S$, $b^j\notin S$, $j=1, 2, \cdots$. Form right ideals $B_i=\{x\mid b^ix\in S\}$, $i=1, 2, \cdots$; since R has the ascending chain condition on right ideals, there exists a positive integer n such that $B_n=B_{n+1}$.

Consider now the right ideal $A=\{b^nr\mid r\in R\}$. Let (S, A) and (S, a) be the ideals generated by $S\cup A$ and $S\cup\{a\}$ respectively. $S\subset(S, a)$ since $a\notin S$, and $S\subset(S, A)$ since $b^{n+1}\in A$ and $b^{n+1}\notin S$. We next show that $S=(S, a)\cap(S, A)$. We already have $S\subseteq(S, a)\cap(S, A)$; suppose $x\in(S, a)\cap(S, A)$. Now $x\in(S, a)$ implies that x can be written in the form $x=\sigma+ar+am$, where $\sigma\in S$, $r\in R$, and m is an integer. Thus $bx=b\sigma+bar+bam$ is in S. Similarly $x\in(S, A)$ means that we can express x in the form $x=\sigma'+b^nr$ and in turn we obtain $bx=b\sigma'+b^{n+1}r$, $b^{n+1}r\in S$, $r\in B_{n+1}=B_n$. Therefore $b^nr\in S$ and $x\in S$. Hence S is reducible, contrary to the original assumption. We conclude that every left divisor of zero in R/S is nilpotent.

Also solved by the proposer.

Metrizable Spaces

5279 [1965, 324]. *Proposed by S. M. Robinson, Union College*

Let X be a completely regular space, X^* its one-point compactification, and $C[X]$ its ring of continuous real-valued functions. For each compact subset F of X, consider the ideal $0_F = \{f \in C[X] : f$ is constantly zero in the neighborhood of $F\}$. Show that X^* is metrizable if and only if there exists a sequence $\{f_n: n=1, 2, \cdots\}$ contained in $C[X]$, such that each ideal 0_F is generated by a subset of $\{f_n\}$ and each function f_n belongs to no free ideal of $C[X]$. (An ideal I in $C[X]$ is free if there exists no point $x \in X$ with the property that $f(x)=0$ for all $f \in I$.)

Solution by the proposer. For any function $h \in C[X]$, let $Z_h = \{x \in X : h(x)=0\}$, and let us assume first that X^* is a compact metric space. Then X is a locally compact separable metric space, and we may assume it has a countable base $\{U_i : i=1, 2, \cdots\}$, where cl U_i is compact for each i. Let $h_n(x) = d(x, \text{cl } U_n)$, and let $\{f_n: n=1, 2, \cdots\}$ be the collection of finite sums of members of $\{h_n\}$. Since Z_{h_n} is compact, so is Z_{f_n}. It follows, therefore, that f_n belongs to no free ideal of $C[X]$. (See Lemma 4.10, Gillman and Jerison, *Rings of Continuous Functions*.) We shall prove that for each compact set F, 0_F is generated by the members of f_n contained in it. If $g \in 0_F$, then $F \subset \text{int } Z_g$; hence, there are sets cl U_{i1}, cl $U_{i2}, \cdots$, cl U_{in} such that $F \subseteq \bigcup_{j=1}^n U_{ij}$, $F \subseteq \bigcup_{j=1}^n \text{cl } U_{ij} \subseteq \text{int } Z_g$. If $f_n = \sum_{j=1}^n h_{ij}$, we observe that $Z_{f_n} \subset \text{int } Z_g$, and this implies that g is a multiple of f_n. (See *Rings of Continuous Functions*, p. 21, 1D.)

The sufficiency of the condition is now established by showing that X is a locally compact separable metric space. For each n, let $U_n = \text{int } Z_{f_n}$, and let $\{V_i : i=1, 2, \cdots\}$ be the family of finite intersections of members of $\{U_n\}$. Now, let U be an arbitrary open set and $x \in U$. Since X is completely regular there is an element g of $C[X]$ such that $x \in \text{int } Z_g \subseteq Z_g \subseteq U$. Now, $\{x\}$ is compact and $g \in 0_{\{x\}}$ so that our hypothesis assures us of functions $m_1, m_2, \cdots, m_n$ in $C[X]$ and members $f_1, f_2, \cdots, f_n$ of the sequence such that $g = \sum_{i=1}^n m_i f_i$, and $f_i \in 0_{\{x\}}$. Now, $\bigcap_{i=1}^n \text{int } Z_{f_i} \subseteq Z_g$, so that we have $x \in V_j \subseteq Z_g \subseteq U$, where $V_j = \bigcap_{i=1}^n U_i$. This proves that the countable collection $\{V_n\}$ is a base, and the proof will be completed by showing that for each $x \in X$, the collection $\{\text{cl } V_n : x \in V_n\}$ is a base of compact neighborhoods of x. In fact, all we need show is that the sets cl V_n are compact. Since f_n belongs to no free ideal of $C[X]$, we see that cl $U_n = Z_{f_n}$ is compact (see *Rings of Continuous Functions*, p. 61, 4E) and this implies that cl V_n is compact.

Sequential Topologies

5299 [1965, 556]. *Proposed by S. Baron, McGill University, Montreal*

Find a set of necessary and sufficient conditions on a topological space $(X, \mathfrak{I})$ such that for any topological space $(Y, \mathfrak{I}')$ and function $f: X \to Y$, f is continuous if and only if $x_n \to x$ in X implies $f(x_n) \to f(x)$ in Y.

Composition of Solutions by Solomon Leader, Rutgers—The State University, and the proposer. Given a topological space $(X, \mathfrak{J})$ the topology $\mathfrak{J}$ induces a topology $\bar{\mathfrak{J}}$, called the *sequential* topology, consisting of all sequentially open sets U: $U \in \bar{\mathfrak{J}}$ if and only if every sequence converging to a point in U is eventually in U. Clearly $\mathfrak{J} \subseteq \bar{\mathfrak{J}}$ and both topologies yield the same convergent sequences. So $\bar{\bar{\mathfrak{J}}} = \bar{\mathfrak{J}}$. $\mathfrak{J}$ is called *sequential* if and only if $\mathfrak{J} = \bar{\mathfrak{J}}$, that is, every sequentially open set is open. We contend that this is a solution to the problem.

(1) Sequential continuity of f: $(X, \mathfrak{J}) \rightarrow (Y, \mathfrak{J}')$ is equivalent to continuity of f: $(X, \bar{\mathfrak{J}}) \rightarrow (Y, \mathfrak{J}')$.

Proof. Given sequential continuity, let $U \in \mathfrak{J}'$, $x \in f^{-1}U$, and $x_n \rightarrow x$. Then $fx \in U$ and $fx_n \rightarrow fx$. Since U is open, $fx \in U$ eventually. That is, $x_n \in U$ eventually. So $f^{-1}U \in \bar{\mathfrak{J}}$. The converse holds because continuity implies sequential continuity, and sequential continuity is the same for both $\mathfrak{J}$ and $\bar{\mathfrak{J}}$.

(2) σ and $\mathfrak{J}$ are identical topologies in X if and only if they admit the same continuous mappings on X.

Proof. Given the latter condition, the identity on (X, σ) into $(X, \mathfrak{J})$ is continuous since it is continuous on $(X, \mathfrak{J})$ into $(X, \mathfrak{J})$. Hence, $\mathfrak{J} \subseteq \sigma$. Similarly, $\sigma \subseteq \mathfrak{J}$. (The converse is trivial.)

Now from (1) and (2) we get the equivalence of:

(i) Continuity on $(X, \mathfrak{J})$ is equivalent to sequential continuity.

(ii) $\bar{\mathfrak{J}}$ and $\mathfrak{J}$ admit the same continuous mappings on X.

(iii) $\bar{\mathfrak{J}} = \mathfrak{J}$. ($\mathfrak{J}$ is sequential.)

REMARKS. Every first countable space is sequential, but the converse is false, a counterexample being the one-point compactification of an uncountable discrete space. S. P. Franklin, *Spaces in which sequences suffice*, Fund. Math., 57 (1965) 107–115, proves that (iii) is equivalent to each of the following:

(iv) X is the quotient of a first countable space.

(v) X is the quotient of a metric space.

(See also R. M. Dudley, *On sequential convergence*, Trans. Amer. Math. Soc. 112 (1964), 483–507.)

Also solved by H. J. Biesterfeldt, Jr.

Differences of Square-Free Integers

5300 [1965, 673]. *Proposed by A. M. Vaidya, Texas Technological College*

Prove that every positive integer can be expressed as a difference of two square-free integers. Compare E 1627 [1964, 686].

I. *Solution by D. L. Silverman, Hughes Aircraft Company, Los Angeles, Cal.* If a positive integer n existed such that every pair of positive integers $(k, n+k)$ contained at least one member with a square factor, then the relative frequency of square-free integers would be asymptotic, if at all, to some number $r \leqq \frac{1}{2}$, in contradiction to the well-known value $6/\pi^2$. (See, e.g., Niven and Zuckerman, *An Introduction to the Theory of Numbers*, p. 226. Th. 11.5.)

II. *Solution by Bob Prielipp, University of Wisconsin.* Let k be a positive integer. Then by Advanced Problem 5286 [1965, 429], there exist infinitely many primes p for which $p+k$ is square-free. Since $k=(p+k)-p$ and p is obviously square-free, we have the desired result, moreover, in infinitely many ways.

Also solved by W. R. Becker, M. G. Greening (Australia), R. L. McFarland, Ivan Niven, C. B. A. Peck, Charles Ryavec, E. G. Straus, K. S. Williams (England), and the proposer.

Editorial Note. C. B. A. Peck has observed the results of the problem in W. Sierpinski, *Elementary Theory of Numbers*, Warsaw, 1964, p. 34. See also Sierpinski's note, *Sur une propriété des nombres naturels*, Revue de Mathématiques Élémentaires, v. 19/2, 1964, 27–28. The following theorem is found in this note:

Let E be a set of positive integers with upper density $>\frac{1}{2}$. Then every integer may be represented as a difference of elements in E in infinitely many ways.

The note also contains a reference for the following generalization given in 1922 by M. T. Nagell:

If the elements a_i, b_i of the number pair (a_i, b_i) are not simultaneously divisible by the square of a prime number, $i=1, 2, 3$, then a_1x+b_1, a_2x+b_2, a_3x+b_3 are all square free for infinitely many positive integers x.

The results above now correspond to the special case $a_1=a_2=1$, $b_1=0$, b_2 an arbitrary integer.

Normal Subgroups of a Finite Group

5301 [1965, 673]. *Proposed by Hermann Simon, McGill University*

Let H be a subgroup of the finite group G. Let q be a prime which divides the order of G but does not divide the index of H in G. Show that H is an invariant subgroup if and only if the index of the normalizer of H does not exceed q and, for every $x\in G$, $H\cap(x^{-1}Hx)$ is either $\{1\}$ or H.

I. *Solution by E. M. Dieckmann, Washington University, St. Louis, Mo.* If H is normal in G, the conditions are clearly satisfied. Conversely, suppose that the stated conditions are satisfied. It follows that the index $|G\colon N(H)|$ of $N(H)$, the normalizer of H, is strictly less than q because $|G\colon H|=|G\colon N(H)|$ $|N(H)\colon H|$, and hence for each $y\in G$ there exists an $r<q$ such that y^r is in $N(H)$. In particular, let y be an element of H of order q (since q divides the order of H), let $x\in G$, and choose $r<q$ so that $(x^{-1}yx)^r$ is in $N(H)$. Since $x^{-1}yx$ is of order q and $(q, r)=1$, it follows that $x^{-1}yx$ is in $N(H)$. But then $x^{-1}yx$ is in H, since otherwise the group $N(H)/H$ would have an element of order q, which is impossible since q does not divide $|N(H)\colon H|$. Hence $x^{-1}Hx\cap H=\{1\}$ so that $x^{-1}Hx\cap H=H$ or $H=x^{-1}Hx$. This shows that H is normal in G.

II. *Solution by J. J. Martinez, University of Pittsburgh.* If H is invariant the two conditions are immediate, so we assume the conditions and prove H invariant.

Let $t\leqq q$ be the index of the normalizer of H and so also the number of distinct conjugates of H in G. Sq will represent a q-Sylow subgroup of G, and there exists at least one Sq in H. Each of the q-Sylow subgroups of G, since they are all conjugate in G, is in a conjugate of H.

If H is not invariant, $1<t\leqq q$, and q cannot divide $t-1$. But by Sylow's

theorem and because no Sq is contained in distinct conjugates of H, $(1+kq)t=1+k'q$, which gives $t-1=q(k'-tk)$, a contradiction.

Also solved by D. M. Bloom, J. R. Durbin, E. W. Ewing, C. C. Lindner, M. D. Mavinkurve (India), C. E. Olson, C. S. Queen, Preston Stein, C. P. Seguin, Gomer Thomas, and the proposer.

Non-complementable Manifolds

5302 [1965, 674]. *Proposed by Pat Garrett, New Mexico State University*

Show that the null manifold of a compact linear operator on a Banach space need not be complemented, i.e., need not be the image of a continuous projection on the domain.

I. *Solution by L. J. Wallen, Stevens Institute of Technology.* If $\phi(x)\in L_1[0, 2\pi)$, define the linear operator T by

$$T\phi(x) = \sum_{n=-\infty}^{-1} \frac{a_n}{|n|^3} e^{inx} \qquad \text{if } \phi(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$

$T\phi=0$ if and only if $a_n=0$, $n=-1, -2, \cdots$, i.e., $\phi\in H_1$. T is compact, for if $f^{(k)}\sim\sum_{-\infty}^{\infty} a_n^{(k)}e^{inx}$ and $\|f^{(k)}\|_1\leqq M$, then $Tf^{(k)}\in C_1$ and

$$\left\| \frac{d}{dx} Tf^{(k)} \right\|_\infty \leqq M \sum_1^\infty 1/n^2.$$

But it is a well-known result of D. J. Newman that H_1 is not complemented in L_1. (See K. Hoffman, *Banach Spaces of Analytic Functions.*)

II. *Note by Joseph L. Ercolano, I.B.M., Yorktown Heights, N. Y.* The question was originally posed by Banach in his classic work, *Théorie des Opérations Linéarires*, for spaces L^p and l^p, $1<p\neq 2$. It has been answered in the negative by several writers who have constructed subspaces of l^p which are not topologically complementable. References follow.

1. F. J. Murray, On complementary manifolds and projections in spaces L_p and l_p, Trans. Amer. Math. Soc., 41 (1937) 138–152.
2. A. Sobczyk, Projections in Minkowski and Banach spaces, Duke Math. J., 8 (1941) 78–106.
3. ———, Projections of the space (m) on its subspace (c_0), Bull. Amer. Math. Soc., 47 (1941) 938–947.
4. H. Komatuzaki, Sur les projections dans certains espaces du type (B), Proc. Imp. Acad. Tokyo, 16 (1940) 274–279.
5. ———, Une remarque sur les projections dans certains espaces du type (B), Proc. Imp. Acad. Tokyo, 17 (1941) 238–240.
6. R. S. Phillips, On linear transformations, Trans. Amer. Math. Soc., 48 (1940) 516–541.

Also solved by the proposer.

Mean Value Theorems in E_k

5303 [1965, 674]. *Proposed by H. S. Shapiro, University of Michigan*

Let ρ be a finite measure on the Borel sets of E_k (Euclidean k-space) such that

$\rho(E_k)=1$, and which is rotation-invariant: $\rho(E)=\rho(TE)$ for every Borel set E and every orthogonal transformation T. Show that, if $u(x)$ is harmonic in E_k and $\int|u|\ |d\rho|<\infty$, then $u(0)=\int u(x)d\rho$.

Solution by L. J. Wallen, Stevens Institute of Technology. Let G be a compact group acting continuously and transitively on the compact Hausdorff space X and let dg denote the (normalized) Haar measure on X. Then it follows immediately from the Riesz theorem that there is a unique regular Borel measure on X satisfying

$$\int_G f(gx)dg = \int_X f(\xi)d\mu(\xi) \qquad x\in X$$

and μ is invariant under G. In case $G=O(n)$, the rotation group in E_n, and $X=rS_{n-1}$, the sphere of radius r about 0, then $\mu=\bar{\omega}_r$, the ordinary normalized Euclidean measure on rS_{n-1}.

Now let $\alpha=\int ud\rho$. Then, since ρ is invariant, $\alpha=\int_{E_n}u(R\xi)d\rho(\xi)$ and

$$\begin{aligned}\alpha &= \int_{O(n)}\int_{E_n} u(R\xi)d\rho(\xi)dR = \int_{E_n}\int_{O(n)} u(R\xi)dRd\rho(\xi)\\ &= \int_{E_n}\int_{|\xi|S_{n-1}} u(\eta)d\bar{\omega}_{|\xi|}(\eta)d\rho(\xi) = u(0)\end{aligned}$$

by the usual mean value theorem for harmonic functions.

Also solved by the proposer.

Self Generating Runs

5304 [1965, 674]. *Proposed by William Kolakoski, Carnegie Institute of Technology*

Describe a simple rule for constructing the sequence

$$122112122122112112212112122112112122122112\ \cdots.$$

What is the nth term? Is the sequence periodic?

Solution by Necdet Üçoluk, Clarkson College of Technology. Let $\{x_n\}$ be a sequence not eventually constant. One can consider a sequence of blocks over $\{x_n\}$ by grouping all consecutive equal numbers in the same blocks. The lengths of these blocks will give a sequence $\{\hat{x}_n\}$ associated with $\{x_n\}$. The given sequence $\{a_n\}$ can be defined as follows: $a_1=1$, $a_n=1$ or 2, and $\{a_n\}=\{\hat{a}_n\}$, for all $n=1, 2, \cdots$. So $\{a_n\}$ is constructed uniquely. Indeed, $a_1=1$, so $\hat{a}_1=1$, hence the first block of $\{a_n\}$ is of length only 1, therefore $a_2\neq 1$, so $a_2=2$. Since $\hat{a}_2=a_2=2$, the second block in $\{a_n\}$ will consist of two 2's, so a_3 and then $\hat{a}_3$ equals 2. Since the third block in $\{a_n\}$ must start with 1, it will consist of two 1's, because $\hat{a}_3=2$. Therefore $a_4=a_5=1$. Continuing in this fashion, $\{a_n\}$ is constructed recursively. If $\{s_n\}$ is the sequence of partial sums of $\sum\hat{a}_n$, then

$a_n=\frac{1}{2}(3+(-1)^m)$ with $s_{m-1}<n\leqq s_m$, since consecutive blocks contain different numbers, odd-numbered blocks contain 1's while the others consist of 2's.

$\{a_n\}$ cannot be periodic. If $\{a_n\}$ were a periodic sequence, say after $n=n_0$, with the minimum period N, then $\{\hat{a}_n\}$ would be periodic after n_0 with the period N. The periodicity of $\{\hat{a}_n\}$ would induce another period $N_1>N$ for $\{a_n\}$ after a certain index n_1 (actually $n_0<n_1<2n_0$). But $N<N_1<2N$, since $\{a_n\}$ can have only blocks of length one or two, therefore a segment of $\{\hat{a}_n\}$ having N elements produces a segment in $\{a_n\}$ of length more than N but less than $2N$. This contradicts the fact that N is the minimum period for $\{a_n\}$. But N_1 must also be a multiple of N, and the non-periodicity of $\{a_n\}$ now follows from this contradiction.

Also solved by Walter Bluger, H. Brandt Corstius (Netherlands), Paul Cull, Jack Dix, R. F. Jackson, Norman Miller, Julius Nadas, C. E. Olson, C. B. A. Peck, J. R. Purdy, Donald Quiring, and Judith Richman.

Nadas generalized the given problem by considering sequences as above but using r integers.

Miller observed that the first 42 digits in the given sequence of 1's and 2's is obtained when each vowel is replaced by 1, each consonant by 2 in the following:

Indian and Ethiopian, Syrian, Israeli, Arab, Persian.

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: K. O. MAY, University of California, Berkeley and E. P. VANCE, Oberlin College

Materials intended for review should be sent to K. O. May, Dept. of Math., University of Toronto, Toronto 5, Ont., Canada.

Introductory Real Analysis. By M. E. Munroe. Addison-Wesley, Reading, Mass., 1965. viii+198 pp. $8.50.

Here is a carefully written introduction to real analysis which can be profitably used at the advanced undergraduate level. The book contains seven chapters entitled: 1. Logic; 2. Real Numbers; 3. Limits; 4. Series; 5. Metric Spaces; 6. Uniformity; 7. Calculus. As one might guess by glancing at the chapter headings, such terms as "differentiable function" or "integral" are not considered until Chapter 7. The first six chapters bring the reader to the point where he can handle the concepts of calculus in a rigorous manner.

The book is written with the student in mind. When a difficult concept or theorem is considered, the author takes time to show the student what is going on. For example, in Chapter 6 one finds, among other things, proofs of the Moore-Osgood Theorem and the Stone-Weierstrass Theorem. These proofs are

followed by short intuitive arguments which should help the student understand the idea behind the proof.

The quantifiers $\forall$ and $\exists$ are introduced in Chapter 1 and are used throughout the book. Munroe takes pains to make it clear from the beginning just how these quantifiers are used and, whenever appropriate, stresses the importance of writing the quantifiers in the proper order. A comment about the format is perhaps in order here. Some mathematicians find statements written in logical notation unpleasant to read. At the price of space, many statements of this type are broken into quantified components, each of which is displayed. This display is somewhat more pleasing to the eye.

A word about Chapter 7. This chapter provides an introduction to differential and integral calculus from the modern viewpoint. After a discussion of manifolds, the author defines "differentials" in terms of functions on tangent bundles. He then proceeds to consider integrals, line integrals, multiple integrals and surface integrals, ending with a proof of Stokes' Theorem. As the author distinguishes between "advanced calculus" and "real analysis" and considers this book to be of the latter type, such topics as Fourier Series, Stieltjes integration and the like are not considered at all.

Munroe takes the trouble to use logically correct notation for limits. Whether one finds this an advantage or disadvantage is probably more a matter of taste than anything else. (Some students might find his notation a bit confusing at first. To some extent, the author admits this when in the latter part of the book he frequently reverts to the more standard notation.)

The sections end with a collection of exercises. These exercises contain many counterexamples and should be looked upon as an essential part of the book.

ANDREW BRUCKNER, University of California, Santa Barbara

Functions of Several Variables. By Wendell H. Fleming. Addison-Wesley, Reading, Mass., 1965. 337 pp. $9.75.

This book is a clear and concise treatment of the usual material in an advanced calculus course with a strong emphasis on presenting definitions and proofs in a coordinate-free manner. Noteworthy for what is intended to be an undergraduate course are the following: distinction is maintained between an n-dimensional euclidean space and its dual; the exterior algebra and the calculus of exterior differential forms are used; a fairly complete treatment for the Lebesgue integral (including convergence theorems and the L^p spaces); and manifolds are presented early. The book concludes with an extensive discussion of integration of forms on manifolds, presenting such material in the manner of modern differential geometry. The book omits none of the usual advanced calculus material, such as the implicit function theorem, extrema, Jacobians, etc. Prerequisites include linear algebra and some elementary topology of euclidean n-space; most of this is included in the appendices. The exercises are numerous and suitable for undergraduates.

R. T. HARRIS, Duke University

Languages with Expressions of Infinite Length. By Carol R. Karp. Studies in Logic and the Foundations of Mathematics. North Holland Publishing Co., Amsterdam, 1964. xix+183 pp. $6.75.

This is a research monograph that gives for the first time a connected and detailed report on some of the important work done by the author, Tarski, Hanf and Scott on logics for languages with infinitely long expressions. Such languages arise in a natural way by extending the language of first-order logic in order to be able to express such notions as archimedean order, torsion, well-order, well-foundedness and the like. Among the central problems that arise from such extensions are the following: (a) to give a satisfactory mathematical framework in which to develop the syntax of systems of infinitely long expressions (including in particular a treatment of the notion of substitution into such formulas); (b) to define formal deductive systems for such languages which allow, ideally, to deduce all valid sentences of the language; (c) to investigate whether and how the major results of finitary logic can be generalized.

The main results in this work are centered around these questions. In particular, a rather complete picture is given for question (b) both for infinitary propositional and predicate logics for an extensive class of such languages. Also the cross connections to Boolean Algebra which were so illuminating in the finitary case are made and again are found quite fruitful.

The style of writing is concise but lucid and should make the book readily accessible to a student moderately versed in mathematical logic and modern algebra. Two preliminary sections (on set theory and Boolean Algebra) make an excellent frame of reference.

ERWIN ENGELER, University of Minnesota

Lectures on Elliptic Boundary Value Problems. By Shmuel Agmon. Van Nostrand, Princeton, N. J., 1965. 291 pp. $3.95.

The first part (Secs. 1–9) deals with the existence of solutions to the Dirichlet problem for an elliptic equation of any order, and with differentiability properties of the solutions of elliptic equations. These two different topics are dealt with by the same tools of á priori estimates, weak derivatives, etc. After giving (Secs. 10–12) some results on general coerciveness and some special boundary conditions (other than the Dirichlet conditions), the author proceeds (Secs. 13–16) to solve eigenvalue problems in a setting which is applicable to general boundary conditions, both for self-adjoint and non-self-adjoint elliptic operators. The asymptotic behavior of the eigenvalues and expansion theorems are derived. This treatment of the eigenvalue problem is an original contribution of the author.

Prerequisite: Real and complex variables, and Hilbert space. To be well motivated, the reader should be familiar with second order elliptic equations.

A. FRIEDMAN, Northwestern University

Monte Carlo Methods. By J. M. Hammersley and D. C. Handscomb. (A Methuen Monograph.) Wiley, New York, 1964. 178 pp. $4.75.

This monograph maintains the excellent standards of the Methuen Monograph series, being written by two of the most successful users and developers of the Monte Carlo technique. After a brief introductory chapter and a rapid survey of some probability and statistical theory, the authors turn to a detailed discussion of random number generators and quasirandom sequences. A short and somewhat unsatisfying chapter on simulation follows; references are given to some fifty applications. (The bibliography contains in all about 300 items.) I feel the authors could have placed more emphasis on the possibility of experimentation in simulation studies, and on the statistical problems it raises.

Two pivotal chapters describe the main variance-reducing techniques; several other devices are mentioned in later chapters. The efficiencies of the main "swindles" are compared numerically in the case of a simple one-dimensional integral; it is a pity the same treatment was not afforded a less trivial problem.

In the second half of the book various applications are described in detail. These expositions are uniformly excellent. Not the least valuable aspect of the book is its documentation of the authors' comment that "Monte Carlo methods constitute a fascinating, exacting, and often indispensable craft. . . ." Emphasis on the last word, please.

C. L. MALLOWS, Bell Telephone Laboratories

The Fourier Transform and its Applications. By Ron Bracewell. McGraw-Hill, New York, 1965. 381 pp. $11.75.

As the author points out, this is not a text for a mathematics course on the Fourier integral; it is designed for a course given to electrical engineers at the first year graduate level. The terms function and waveform are used interchangeably, as are the terms Fourier transform and spectrum; few theorems are proved and other important theorems are discussed but not carefully stated even for special cases (the Fourier inversion theorem, the central limit theorem). Many interesting applications of Fourier transforms to electrical engineering problems are given.

The first five chapters are devoted to background material needed for the study of Fourier transforms. Convolution is discussed at length and its importance is properly emphasized throughout the text; Lighthill's development of generalized functions is outlined and waveforms are then taken to be generalized functions in this sense.

In Chapter 6 the basic theorems for operations with Fourier transforms are given; Plancherel's theorem is called Rayleigh's theorem, and Parseval's theorem is called the Power theorem. Chapter 7 and Chapter 8 give some interesting examples of using transforms. Chapter 11 discusses the Laplace transform and some of its applications; in Chapter 12 a special case of the Hankel transform ($\nu=0$), the Mellin transform, the z transform, the Abel transform, and the

Hilbert transform are defined. The remaining chapters are given to applications. The chapter titles are: 9. Electrical waveforms, spectra, and filters; 10. Sampling and series; 13. Antennae; 14. Television image formation; 15. Convolution in statistics; 16. Noise waveforms; 17. Heat conduction and diffusion. Chapters 18 and 19 are tables.

The text is written in a readable style and the typography is excellent. The author states in the introduction that the book started out to be a pictorial dictionary of Fourier transforms. This book must obviously be compared to Zemanian's *Distribution Theory and Transform Analysis*, McGraw-Hill, 1965. Zemanian's book is mathematically well founded and the engineering applications can be understood by any senior mathematics student. Bracewell's book shows no interest in the mathematics *per se* but is very effective in his viewpoints and in his applications. To understand the applications a student needs to know some antenna theory. For the type of course for which this book is designed it should and probably will become a standard text. Mathematically it should be supplemented by something like *Distributions, Complex Variables, and Fourier Transform*, Hans Bremermann, Addison-Wesley, 1965.

T. K. BOEHME, University of California, Santa Barbara

Introduction to p-adic Numbers and Valuation Theory. By George Bachman. Academic Press, New York, 1964. 173 pp. $3.45 (paper), $6.50 (cloth).

The first two chapters of this book are a leisurely introduction to rank one valuation theory and are written for a reader with a minimum of prerequisites. Special attention is paid to the p-adic field Q_p; and p-adic analysis, the structure of Q_p, and the roots of polynomials over Q_p are discussed.

The remaining three chapters go more deeply into the subject and thus require more sophistication on the part of the reader, but a course in algebra through Galois theory should suffice. The pace is still leisurely and many examples and exercises are given to illustrate what is proved. In Chapter III the concepts of place, valuation, and valuation ring are introduced and tied together; a proof of the place extension theorem is given; and an application to the concept of integral closure is presented. In Chapter IV the author digresses to the theory of normed linear spaces in order to prove the Gel'fand-Tornheim theorem (a little spectral theory is given). In Chapter V he proves the standard results concerning the existence and number of extensions of a given valuation. Finally, an appendix is given outlining the results that are needed from algebra.

Throughout this book one senses the influence of E. Artin both in the style and in the presentation. In fact, the author mentions that he based a large part of Chapter III on some notes of Artin's. On the whole the book makes very nice reading, but there is an occasional tendency on the part of the author to introduce unnecessary words and concepts; e.g., he proves the Hahn-Banach theorem for sublinear functionals instead of norms, in a book on valuation theory. Also, on occasion, the statement of a result or a proof could be handled both

more conceptually and more elegantly. Both these faults are only minor intrusions in the book.

The scope of the book is limited (the author does not even mention the connection between the valuations of a number field and the ideal structure of its ring of integers), but the book should prove to be a useful addition to the literature on the subject. It gathers together various topics in valuation theory not easily accessible on an elementary level and offers a fine introduction to the subject.

W. W. ADAMS, University of California, Berkeley

Mathematical articles in the Encyclopaedia Britannica, Chicago, 1965.

Here is a counterexample to the bromides that encyclopaedias contain nothing of interest to mathematicians and that the eleventh edition of the Britannica in 1910 was the last good one. Among some 400 quarto pages devoted to mathematics (including about 100 substantial articles, over 100 biographies, and many shorter entries) any mathematician is likely to find some that interest him, and the current edition reflects contemporary mathematical thinking far better than the eleventh. For this happy state of affairs congratulations are due the present mathematical editor, Eldon Dyer, and his predecessors Irving Segal and Edwin Spanier, who managed over a period of years to secure an impressive number of high quality expositions. Among these we single out the following for combining up-to-date information with insights and historical perspective.

Algebra (A. P. Mattuck); Algebra, History of (O. Ore); Algebraic Geometry (W. Hodge); Algebras, Linear (D. Zelinsky); Analysis (S. Bochner); Analysis, Abstract (G. W. Mackey); Calculus of Variations (C. B. Morrey); Fourier Series (A. Zygmund); Functions, Analytic (E. Hille); Geometry (H. Freudenthal); Groups, Transformation (L. Zippin); Integration and Measure (M. Loève); Lattice Theory (G. Birkhoff); Logic, History of Modern Logic (A. Church); Mathematics, Foundations of (S. C. Kleene); Operators, Theory of (R. V. Kadison); Tensor Algebra (I. M. Singer); Topology, Algebraic (W. S. Massey); Topology, General (R. L. Wilder).

The list could be more than doubled without significantly decreasing either the distinction of the authors or the quality of the exposition. The major articles include several high level technical surveys such as Differential Geometry (S. S. Chern) and Manifolds (P. A. Smith). Generally elementary mathematics fares less well, topics being treated in a rather pedestrian and sometimes old-fashioned way. Besides major articles there are some fine briefer pieces such as Binomial Theorem (L. Carlitz), Metatheory (J. C. Kemeny), and many on logic by A. Church. Most of the short articles, however, especially those on geometry, might better have been incorporated into bigger ones. The biographies are well done (Ore being the biggest contributor), but mathematicians get too little space compared with minor literary, political, and religious figures.

In such an enormous work, which has been subject to "continuous" (annual) revision since 1936, uneven execution is to be expected. The article "Mathe-

matics, Articles on" is a good idea, but out of date. Indexing and cross-referencing are inadequate. For example, "cobordism" appears in two articles but not in the index. There are overlaps and inconsistencies. For example, in his article on Analysis Bochner indicates the importance in topology of the Kronecker integral, but articles on topology do not mention it. No doubt specialists can find faults in articles in their fields, and the non-specialist may notice such weaknesses as the semantic confusions in Binary Numbers and the failure to deal separately with heuristic, graph theory, linear programing, and probability. But one may expect continued annual additions and improvements.

The Britannica deserves a place in mathematical literature. By suggesting appropriate articles, the teacher can put his students in direct communication with leading mathematicians and perhaps give them some notion of the living mathematics that is often so poorly reflected in our arid texts. Graduate students would find the survey articles helpful for orientation or cramming. Mature mathematicians may find useful introductions to unfamiliar fields and interesting commentary on their specialties.

KENNETH O. MAY, University of California and Carleton College

Methods of Real Analysis. By R. R. Goldberg. Blaisdell Publishing Company, Waltham, Mass. 1964. 353 pp. $9.00.

One of the problems in the undergraduate curriculum in many schools has to do with a suitable transition from elementary calculus to a solid course in real variables. If this problem has been intensified in the past by a lack of appropriate textbooks, the situation has now changed. The material in this book is quite simple at the beginning but still sufficiently advanced in the latter chapters to be an excellent transition from elementary calculus to a course in real variables. It contains many of the usual topics of an advanced calculus course, but the treatment is up-to-date throughout. New topics, definitions, and theorems are well motivated before being precisely stated. There is a wealth of well graded exercises as well as illustrative examples. The author writes in a refreshing style. At times it is almost conversational. It should be easy for students to read and to want to read.

The selection of material and organization has been given a lot of thought. Chapter 1 begins very simply with the elementary notions of sets and functions, some properties of the real numbers, and the least upper bound axiom. Chapters 2 and 3 contain a thorough but elementary treatment of sequences including summability of sequences, and the usual material on infinite series. Summation by parts, some material on real numbers, and decimal expansions are topics covered in Chapter 3.

In Chapter 4 the concept of limits and metric spaces are introduced. Chapter 5 is a study of continuous functions on metric spaces and Chapter 6 deals with connectedness, compactness, and completeness. The Baire category theorem is included. In these chapters a thorough study is made of limits and continuity, and the fundamental ideas for many topics in the rest of the book are presented.

Chapter 7 begins with an excellent treatment of the Riemann integral followed by the theory of derivatives, and as one would expect, the fundamental theorems of calculus that link together the concepts of the integral and the derivative. The final topic in this chapter is that of improper integrals. The elementary functions are treated in Chapter 8. They are all defined by integrals and their equivalence to these functions obtained by the more common definitions is established. The usual material on Taylor and Maclaurin series and L'Hospital's rule is also in Chapter 8. The next chapter covers the notions of sequences of functions, uniform convergence, and Abel summability.

In Chapter 10 the metric space $C[a, b]$ of continuous functions on a closed bounded interval is defined and the following theorems about this metric space are proved: (1) The space $C[a, b]$ is complete; (2) The Weierstrass approximation theorem; (3) The Picard existence theorem for differential equations; (4) Arzela's theorem.

In Chapter 11 a careful but readable theory of measure and the Lebesgue integral is developed. The metric space $L^2[a, b]$ is defined and it is proved that the set of continuous functions on $[a, b]$ is dense in $L^2[a, b]$. The Riesz-Fischer theorem is proved, and Lebesgue integrals on $(-\infty, \infty)$ and in the plane are treated briefly.

Chapter 12 contains a brief development of the elements of Fourier series including the L^2 theory of Fourier series and orthonormal expansions in $L^2[a, b]$.

The author of this book has attempted to write an introductory text on the theory of functions of real variables for a course that will replace the traditional course in advanced calculus. This goal has been accomplished. Some who are inspecting the book with the intention of possibly adopting it for their courses might criticize the fact that there is no multi-dimensional calculus included, but a thorough knowledge of the material that is included will prepare a student well for advanced courses in analysis.

RALPH L. SHIVELY,
Oak Ridge National Laboratory, and Lake Forest College

Oeuvres de Camille Jordan, 4 vols. 1961–1964. Tome I, 498 pp., 1961; II, 556 pp., 1961; III, 554 pp., 1962; IV, 610 pp., 1964. Gauthier-Villars et Cie, Paris.

After groping developments in the theory of groups by Ruffin, Lagrange, Cauchy and others, Abel and particularly Galois clarified the basic importance of group theory in the problem of solving algebraic equations by means of expressions involving a finite number of rational operations and extraction of radicals performed on the coefficients of the equation. After Abel's proof of the impossibility of solving algebraically the general irreducible equation of any prime degree $\geqq 5$, further progress was blocked by lack of knowledge of the structure of groups connected by Galois's research with the general irreducible equation of any given degree n.

Present-day mathematicians, excepting the French, may easily think of Camille Jordan (1838–1922) essentially in connection with his important work in this direction. For this there would be considerable justification since in the *Oeuvres*, the first two of the four large volumes are devoted to these topics. The remaining two volumes contain in over 1000 pages, his important contributions to many branches of mathematics (Jordan curve, J. integral, J. content, etc.).

His famous *Traité des Substitutions*, which has been available in many editions and translations, and which was indispensable for algebraists through the last third of the 19th century, and his well-known *Cours d'Analyse* are not included in the *Oeuvres* which contain, in about 130 articles, all the other mathematical work of Jordan, written between 1861 and 1922.

The outstanding importance of Jordan's work in algebra and the theory of groups is characterized by eulogies by Gaston Julia, Jean Dieudonné, Em. Picard and Henri Lebesgue, short passages from which are translated below.

Julia, vol. I, Preface: "It is well known that, after Galois, Jordan was the first and for a long time almost the only one to add basic new results to those bequeathed us by Galois.... For a long period Jordan worked in what amounted to practically complete isolation. Rare indeed were those able to appreciate the true value of his work. Today his articles seem more actual than when they were composed; we see them in their true light and in their fundamental scope, and Jordan now appears to us, together with Galois and Sophus Lie, as one of the three great creators of general group-theory." It should be remembered that these words were written a few years after the end of the first world war, and an objective appreciation of German mathematicians can hardly be expected. This prejudicial attitude shows clearly in some other places of the various orations included in the *Oeuvres*. It may be mentioned that Jordan lost in battle in the years 1914–16 three sons and his eldest grandson.

Dieudonné, vol. I, p. XVII: "At the time Jordan commenced publishing (from 1861 on, to 1922) group theory was still in its infancy; indeed, it only acquired the status of an autonomous discipline with the publication of the *Traité des Substitutions* in 1870."

Lebesgue, vol. IV, p. XX: "In his determination of groups of motions which were applied immediately by students of crystallography, we find . . . the first study of a group of transformations. Here he also introduces the basic notion of the fundamental region of a group."

Picard, vol. IV, p. VIII: "However, it is chiefly in the theory of substitutions and algebraic equations that Jordan has left his deepest imprint. In his *Traité des Substitutions* he has made a profound study of the ideas of Galois, adding fundamental results on primitive groups, transitive groups and composite groups, leading in particular to the composition-factors of a group. These studies permitted Jordan to solve the problem, proposed by Abel, of determining the equations of a given degree which are solvable by radicals (in the classical sense), and of recognizing whether a given equation does or does not belong to this class."

For over-all orientation in the four volumes one may suggest the following:

vol. I: Julia, *Préface*, pp. V–VII.
Dieudonné, pp. XVII–XLII, *Notes relatives à la théorie des groupes finis.*

vol. III: Dieudonné, pp. V–XX, *Notes relatives à l'algèbre linéaire et multilinéaire et à la théorie des nombres.*

vol. IV: Julia, *Préface*, pp. V–VI.
Lebesgue, *Notice sur la vie et les travaux de Camille Jordan*, X–XXIX (particularly XIX–XXVI).
C. Jordan, *Notice sur les travaux de M. Camille Jordan* (1881). *Rédigé par M. C. Jordan à l'appui de sa candidature à l'Académie des Sciences.* pp. 553–581.

Here Jordan gives his own analysis of his mathematical work up to the time of his admission to the Académie in 1881 (age 43), and outlines what amounts to his life's work in mathematics. At the Mathematical Congress after the first world war at Strasbourg in 1920 (not counted as International Congress because mathematicians from enemy countries were excluded), Jordan was by acclamation elected Honorary President.

The mathematical articles in the "*Oeuvres*" are "*reproductions photomécaniques*" of the original papers.

Edited by Gaston Julia, Jean Dieudonné and René Garnier, these important volumes represent a debt of honor paid to an outstanding mathematician and a great man.

A. J. KEMPNER, Boulder, Colorado

Theory of Functions as Applied to Engineering Problems. Edited by R. Rothe, F. Ollendorff and K. Pohlhausen. Dover, New York, 1961. 189 pp. $1.35 (paper).

This work was first published in Germany in 1931 by Julius Springer, Berlin, under the title, *Functionentheorie und ihre Anwendung in der Technik.* This new Dover edition is an unabridged and corrected republication of the English translation published by the Massachusetts Institute of Technology in 1933. The book is based on a series of lectures delivered at the Berlin Institute of Technology, in its Winter Session of 1929–1930. The purpose of these lectures was to acquaint a group of electrical engineers with the fundamentals of the theory of functions and with some of the technical applications of the theory.

The first part of the volume consists of eight introductory lectures by Dr. Rothe. The second part consists of five lectures delivered by five eminent theoretical engineers. The mathematical introduction is a model of brevity and clarity. Starting with the fundamental definitions and concepts of complex quantities, Dr. Rothe develops the concepts of the line integral, integration in the complex plane, Laurent series, residue theorems, the use of function theory in the solution of linear differential equations with constant coefficients, and conformal transformations including the Schwarz transformation.

The second part of the book is devoted to applications of the general theory of functions to problems of primary interest to electrical engineers such as the construction of two-dimensional electric or magnetic fields by means of source-line potentials, the determination of the electric field at the edge of a plane capacitor, the determination of the magnetic field of a slot, the solution of electrical and thermal transient problems, the spreading of electro-magnetic waves along the earth, and the solution of two-dimensional flow problems involving ideal incompressible fluids.

In the opinion of the reviewer, this book is still one of the best short accounts of the theory of functions and its application to two-dimensional field problems despite the fact that it was first published in 1931. With minor changes in notation, the section on transient problems can easily be translated into the modern Laplace transform language and it gives the student several clear examples of the computation of inverse Laplace transforms by the use of contour integrals involving branch-points and branch-cuts.

L. A. PIPES, U.C.L.A.

Basic Concepts of Geometry. By Walter Prenowitz and Meyer Jordan. Blaisdell, Waltham, Mass. 1965. 373 pp. $7.50.

The main objective here is to present various geometrical systems, to show their interrelations, to characterize them according to their respective postulate systems, and, finally, to offer a treatment of Euclidean geometry which meets current standards of rigor. This objective is neatly and effectively accomplished through the basic unifying concept of Incidence Geometry, because each of the well-known geometries (except spherical geometry) is an incidence geometry which is further characterized by some additional postulates concerning parallelism, order, congruence, continuity. Following the introduction of six incidence postulates, about twenty-five models of the theory of incidence are exhibited. Among these are both finite and infinite models, models to satisfy the parallel postulates of Euclidean, Lobachevskian, and Riemannian geometries, projective and affine geometries, and algebraic models.

Part I is designed as preparation to make the reader aware of the necessity to bridge the gap between the classical and modern treatments of Euclidean geometry. Part II contains the abstract development. Although the treatment of the geometry is formal, the entertaining lively style with which the logical principles are discussed adds much in the way of pedagogical advantage. The exercises are ample and carefully selected. This stimulating book should be required reading for all teachers of geometry. The authors do an excellent job of making Euclidean geometry respectable, although they do not propose their approach for use in the high school. It is one thing to prepare the teacher properly for teaching geometry, and quite another to select the appropriate subject matter and standard of rigor for a given level of students.

C. E. SPRINGER, University of Oklahoma

Functions of a Complex Variable and Some of their Applications. By B. A. Fuchs and B. V. Shabat. Pergamon Press, New York, 1964. xvi+431 pp. $10.00.

This is a revised and expanded version of an English translation of the 1959 second edition in Russian. The book is written for Science and Engineering students whose background in vector and real analysis includes such material as line integrals, Green's Theorem and advanced calculus material on real series. There is a short introduction to complex numbers. The approach to complex analysis is through mappings, multiple valued functions and Riemann surfaces, and conformal mapping appears as early as the second chapter.

The use of examples and references is basic to the author's style of presentation of the theoretical subjects. In most cases he gives the basic theorems, mathematical examples, references to more extensive coverage and closes with nontrivial applications. Simple or straight forward proofs are given and the more complicated proofs are outlined with references for the complete proof: e.g. for Cauchy's Theorem, the partial derivatives are assumed continuous. The extensive theoretical work is in series and analytic continuation with a good discussion of natural boundaries, singular points and the point at infinity. A lengthy chapter on the theory of residues includes brief discussions of Mittag Leffler's theorem and infinite products.

At least one half of the book is given to good examples and applications to science and engineering problems. An entire chapter is devoted to stationary plane vector fields and potential theory and most of another to examples of conformal mapping. The chapter on mapping of polygonal domains consists primarily of applications of the reflection principle and the Schwarz-Christoffel integral. There are also brief discussions of Dirichlet's problem, elliptic integrals and elliptic functions and the Laplace transform. The inclusion of references for further study, an index, problems, and a complete set of answers make this a good text for the appropriate student.

A. H. CAYFORD, University of British Columbia

Elements of Mathematical Logic. By P. S. Novikov. Translated from the Russian by Leo F. Boron. Preface and notes by R. L. Goodstein. Addison-Wesley, Reading, Mass., and Oliver and Boyd, London, 1964. xi+296 pp. $7.95.

According to the author's foreword, this book is designed as a text which will be less concise than Hilbert and Ackermann's *Principles* and more elementary than Kleene's *Metamathematics.* The first five chapters succeed very well in fulfilling this goal. Standard material in the propositional calculus and in the first order predicate calculus is covered, and proofs are given clearly and in detail.

In Chapter I the propositional calculus is introduced using truth tables. This is followed by its axiomatization in Chapter II; this chapter includes an interesting section on the independence of the eleven axioms. In Chapter III first order predicate logic is introduced via informal set theoretic interpretations.

The decidability of predicate logics containing only predicates of one variable is demonstrated, and the chapter ends with a proof of the Löwenheim Theorem using Skolem functions. The predicate calculus is axiomatized in Chapter IV, and Skolem normal forms and Maltsev's Compactness Theorem (for countable sets of formulae) are used in the proof of Gödel's Completeness Theorem. In Chapter V axioms for formal arithmetic are given. The notion of a recursive term in formal arithmetic is introduced and used to define primitive recursive functions. General recursive functions are defined, but no theorems about them are given, nor are partial recursive functions mentioned.

Chapter VI gives an interesting new finitary proof of the consistency of formal arithmetic without the axiom schema of induction and also, using the same methods, a proof that there is no consistent finitely axiomatizable extension of this restricted arithmetic in which the schema of induction is provable. Students will probably find this final chapter considerably more difficult than most of the earlier material. An introduction of twenty pages, largely concerned with a description of finitary methods and their role in foundations, is perhaps best read in conjunction with this chapter.

In spite of the book's general clarity, there are occasional lapses. For example, on page 114 the unsolvability of the full predicate calculus is linked with the Fermat problem in a puzzling fashion, and on page 220, in a discussion of primitive recursive functions, it appears that members of the range are chosen arbitrarily when in fact the results of the preceding section are to be utilized.

Other than an occasional proof left as an exercise, there are no problems, nor is there a bibliography.

This book should be a good text for courses designed to give students a good technical background for more advanced studies in logic. With the exception of arithmetic, however, no attempt is made to show connections between logic and other branches of mathematics. Because of this, many persons will prefer a text with greater emphasis on applications to other areas of mathematics, but even these people will find this book an excellent reference for material either omitted from or treated briefly in other books or in class.

PAUL YOUNG, Purdue University

Integral Equations and Partial Differential Equations. By V. I. Smirnov, Addison-Wesley, Reading, Mass., 1964. 811 pages, $17.50.

This book, Volume Four of a course of Higher Mathematics, was translated from the Russian by D. E. Brown and edited by Prof. I. N. Sneddon of the University of Glasgow. The subject matter of this volume is the basic theory of classical mathematical physics.

Chapter 1, entitled *Integral Equations*, is devoted to the study of orthogonal systems of functions and the general theory of Fredholm equations. It contains almost two hundred pages and covers the subject matter well, including the classical theorems and various classes of kernels.

Chapter 2, entitled *The Calculus of Variations*, is the shortest of the four

chapters, covering the subject matter in approximately one hundred pages and containing quite a few examples and applications.

Chapter 3, entitled *Fundamental Theory of Partial Differential Equations*, is divided into three parts. The first part contains the theory of equations of first order, both linear and non-linear, and the associated theory of characteristics, and includes the Cauchy-Kowalewski theorem. The second part is concerned with the fundamentals of second order linear equations. The third part deals briefly with systems of quasi-linear first order equations and second order systems, with several examples originating from physics.

Chapter 4, entitled *Boundary Value Problems*, is the largest of the chapters and is divided into three parts. The first part deals with boundary value problems for ordinary differential equations with the accompanying theory of the Green's function and eigenvalue theory. The second part deals with elliptic equations and the related Dirichlet and von Neumann Problems, with emphasis on the Laplace equation. The third part deals with hyperbolic and parabolic equations.

It is difficult to describe in a limited space the vast amount of material covered in this excellent text. The author goes to great lengths in trying to motivate the reader. The theorems are nicely stated and the proofs are done very clearly. Furthermore, the author supplies quite a bit of background and his examples are well chosen and clearly done.

The organization of the volume is a weak point. The author often refers to material in his previous volumes to start his development rather than giving general bibliographical references. This presupposes possession of the previous volumes. No organized bibliography is given, and the index is incomplete.

MURRAY WACHMAN, General Electric Co.

Equations of the Mixed Type. By A. V. Bitsadze. Translated by P. Zador. Translation edited by I. N. Sneddon. Pergamon Press, New York, 1964. xiii+160 pp. $8.50.

Partial differential equations are classified into various types according to the values of the coefficients. The problems which are appropriate and the general behavior of the solutions depend upon the type. When the equation belongs to different types in different parts of the domain, the equation is of mixed type. The first fundamental studies of equations of mixed type were made by Francesco Tricomi in 1923. It has become an increasingly active area of research, especially in the last twenty years.

The book is a translation from the Russian original of 1959. It is a self contained study, for various mixed equations, of questions such as existence, uniqueness, and construction of solutions. Conditions on the coefficients and boundary curves which would be sufficient for the correctness of particular problems associated with an equation are also studied. A helpful feature is the discussion of generalizations and open questions at the end of each chapter.

After a discussion of canonical forms in the first chapter, the next four chapters study specific equations and problems. Chapter II takes up second order hyperbolic equations with initial data given along the lines of parabolicity. Chapter III considers elliptic equations with a boundary curve of parabolic degeneracy. Chapters IV and V are on the Tricomi problem and other mixed problems.

The references to the bibliography are very helpful. There is also a short index.

A. M. WHITE, Harvey Mudd College

Définition des Fonctions Eulériennes par des Équations Fonctionnelles. By Jean Anastassiadis. Gauthier-Villars, Paris, 1964. 78 pp. 16 F.

This book inevitably invites comparison with Artin's well-known monograph on the Gamma-function; unfortunately, while containing a good deal of information, it is neither as smoothly written nor as well motivated as Artin's book. Examples: $\Gamma(1/2)=\sqrt{\pi}$ is used before it is proved. The function $\mu(x)=\log\Gamma(x)-(x-1/2)\log x+x-1/2\log 2\pi$ is introduced without any motivation and at a time when the $1/2\log 2\pi$ particularly seems a whim of the author. After a detailed proof that Euler's constant $C=\lim_{n\to\infty}(\sum_{k\leqq n}1/k-\log n)$ exists, a nonobvious integral expression for C is quoted from Whittaker and Watson for future use, but without any indication of proof. The author defines the Gamma-function by Gauss' formula

$$\Gamma(x)=\lim_{n\to\infty}\frac{n!n^x}{x(x+1)\cdots(x+n)},$$

and convergence is proved; later he finds it necessary to introduce *ad hoc* and prove separately the existence of the Euler integral. (Starting with the Euler integral instead as definition, Gauss' formula follows from Artin's proof of the Bohr-Mollerup Theorem).

The material which is novel as compared with Artin is primarily due to the author and generally consists in theorems and characterizations in which the Bohr-Mollerup condition "$F(x)$ is logarithmically convex for $x\geqq x_0>0$" is replaced by the (non-equivalent) condition, "$(e/x)^xF(x)$ is nondecreasing for $x\geqq x_1>0$" or a similar one.

The first chapter, in addition to standard material on convex and logarithmically convex functions, contains two tests for logarithmic convexity due to Montel, and the definitions and basic properties of semi-monotone and semi-convex functions. An appendix at the end of the book gives without proof some results of Krull on solutions of $g(x+1)-g(x)=\phi(x)$.

Despite the inelegancies of style, this monograph might well be read with profit as "enrichment" by an undergraduate who possessed the requisite knowledge of French.

S. L. SEGAL, University of Rochester

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

The following mathematicians have been elected to membership in the National Academy of Sciences: Professor A. M. Gleason, Harvard University; Professor Irving Kaplansky, University of Chicago; and Dr. S. M. Ulam, Los Alamos Scientific Laboratory.

Professor Emeritus Richard Courant, New York University, has been named a member of the Academy of Sciences of the U.S.S.R.

Professor Ralph Donnell, Union University, represented the Association at the inauguration of J. D. Alexander, Jr., as President of Southwestern at Memphis on May 3, 1966.

Professor C. M. Lindsay, Coe College, represented the Association at the Dedication of the Science-Classroom Building, West Residence Hall, and Maintenance Center of Clarke College on March 19, 1966.

Professor J. E. Maxfield, University of Florida, represented the Association at the inauguration of J. E. Champion as President of Florida State University on March 15, 1966.

Queen's University: Associate Professor A. H. Lightstone, Carleton College, has been appointed Associate Professor; Professor H. A. Still, Alfred University, has been appointed Associate Professor; Dr. Norman Rice, California Institute of Technology, has been appointed Assistant Professor.

Assistant Professor E. N. Howell, Wisconsin State University, River Falls, has been promoted to Associate Professor.

Assistant Professor E. M. Hughes, Chadron State College, has been promoted to Associate Professor.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

MATHEMATICS IN BIOLOGY

A task committee representing the Committee on the Undergraduate Program in Mathematics (CUPM) and the Committee on Undergraduate Education in the Biological Sciences (CUEBS) is engaged in an effort to collect examples of meaningful applications of mathematics at all levels in the solution of biological problems. This effort is stimulated by the recognition that the coming years will see a striking increase in the quantity and quality of mathematics taught to and used by biologists. The immediate objective of this effort is to place before the scientific community a collection of examples which (1) will add meaning and motivation to mathematics courses attended by biologists and (2) will be cross-referenced according to the field of biology involved for use in biology instruction. Among the long-range objectives is the stimulation of two-way research interchanges between mathematicians and biologists.

Members of the mathematical community who are interested in being kept abreast of these developments are requested to write to the Committee on the Undergraduate Program in Mathematics, P. O. Box 1024, Berkeley, California 94701. Names of other persons who are interested in this information will be welcomed.

SUMMER DUES PAYMENTS

In common with all other organizations charging dues on an annual basis, MAA is faced each year with a great overload of work at the end of the year when nearly all members try to pay dues simultaneously.

It would help spread the work of the Buffalo office if some members would pay dues in advance. May we suggest, therefore, that some (but not all) plan to pay 1967 dues during the present summer. Member's annual dues are still $6. Payments should be sent to: Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, N. Y. 14214.

CALENDAR OF FUTURE MEETINGS

Forty-seventh Summer Meeting, Rutgers, The State University, New Brunswick, New Jersey, August 29–September 1, 1966.

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN
ILLINOIS
INDIANA
IOWA
KANSAS
KENTUCKY
LOUISIANA-MISSISSIPPI, Jung Hotel, New Orleans, La., March 4–5, 1967.
MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA
METROPOLITAN NEW YORK
MICHIGAN
MINNESOTA
MISSOURI
NEBRASKA, University of South Dakota, Vermillion, May 6, 1967.
NEW JERSEY, Rutgers, The State University, New Brunswick, November 12, 1966.
NORTHEASTERN, Trinity College, Hartford, Conn., November 26, 1966.
NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.
OHIO
OKLAHOMA-ARKANSAS, Northeastern State College, Tahlequah, Okla., March/April, 1967.
PACIFIC NORTHWEST
PHILADELPHIA, Villanova University, Villanova, Pa., November 19, 1966.
ROCKY MOUNTAIN
SOUTHEASTERN, Florida Presbyterian College, St. Petersburg, March 31–April 1, 1967.
SOUTHERN CALIFORNIA, San Diego State College, San Diego, March 18, 1967.
SOUTHWESTERN
TEXAS, Austin College, Sherman, April 14–15, 1967.
UPPER NEW YORK STATE
WISCONSIN

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, Rutgers, The State University, New Brunswick, N. J., August 30–September 2, 1966.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Michigan State University, June 19–23, 1967.

ASSOCIATION FOR COMPUTING MACHINERY, Ambassador Hotel, Los Angeles, August 30–September 1, 1966.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

INSTITUTE OF MATHEMATICAL STATISTICS, Rutgers, The State University, New Brunswick, N. J., August 30–September 2, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Houston, Texas, January 28, 1967.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Jack Tar Hotel, Durham, N. C., October 17–19, 1966.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, State University of New York, Stony Brook, September 12–14, 1966.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 7

CONTENTS

AUGUST–SEPTEMBER 1966

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. Ficken, Department of Mathematics, New York University, New York, N. Y. 10453. *Manuscripts of papers should be submitted to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.*

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

THIS IS THE OFFICIAL JOURNAL OF THE

MATHEMATICAL ASSOCIATION OF AMERICA, INC.

(Devoted to the Interests of Collegiate Mathematics)

General Office: SUNY at Buffalo, Buffalo, N. Y. 14214. *Executive Director:* H. M. GEHMAN

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

HIGHER CURVATURES OF CURVES IN EUCLIDEAN SPACE

HERMAN GLUCK, Harvard University

1. Introduction. It is the object of this note to point out a simple algorithm, based on the Gram-Schmidt orthonormalization process, for computing the curvatures of a curve in Euclidean n-space. This algorithm has a number of agreeable features. There is just one formula involved for all the curvatures, and it is very simple. There is no duplication of effort, in the sense that if one calculates the Frenet frame associated with a curve at a given point, then the various curvatures can be obtained immediately from the by-products of this calculation. Finally, the entire procedure is no more involved if the parametrization is not by arc length.

2. The Frenet frame and the curvatures. Let I be an interval in R^1 and $F: I \to R^n$ a C^k-parametrization by arc length. This means that the arc length along the curve from $F(s_1)$ to $F(s_2)$ is $|s_1 - s_2|$, or equivalently that $|F'(s)| = 1$ for all $s \in I$. Suppose that for each $s \in I$, the vectors

$$F'(s),\ F''(s),\ \cdots,\ F^{[r]}(s) \qquad r < k$$

are linearly independent. Applying the Gram-Schmidt orthonormalization process to these vectors, one obtains an orthonormal r-tuple of vectors,

$$(V_1(s),\ V_2(s),\ \cdots,\ V_r(s)),$$

called the *Frenet r-frame* associated with the curve at the point $F(s)$. $V_i(s)$ is easily seen to be of class C^{k-i}.

Formulas for the derivatives $V_i'(s)$ can be obtained by first differentiating the orthonormality relations $V_i(s) \cdot V_j(s) = \delta_{ij}$, to get

$$V_i'(s) \cdot V_j(s) = - V_j'(s) \cdot V_i(s).$$

Combining this with the fact that, for $1 \leqq i \leqq r-1$, $V_i'(s)$ is a linear combination of $V_1(s)$, $V_2(s)$, $\cdots$, $V_{i+1}(s)$, one concludes that

$$V_i'(s) \cdot V_j(s) = 0,$$

except possibly for $j = i-1$ and $j = i+1$. The derivative formulas can therefore be written

$$\begin{aligned} V_1'(s) &= k_1(s) V_2(s) \\ V_i'(s) &= - k_{i-1}(s) V_{i-1}(s) + k_i(s) V_{i+1}(s) \qquad 2 \leqq i \leqq r-1. \end{aligned}$$

There is a bit of a problem with $V_r'(s)$, since there may be no $V_{r+1}(s)$. Given $s_0 \in I$, if $F^{[r+1]}(s_0)$ is linearly independent of $F'(s_0)$, $F''(s_0)$, $\cdots$, $F^{[r]}(s_0)$, then this will also be true in some neighborhood of s_0 in I. For s in such a neighborhood, $V_{r+1}(s)$ can be defined as above and one will have

$$V_r'(s) = - k_{r-1}(s) V_{r-1}(s) + k_r(s) V_{r+1}(s).$$

If $F^{[r+1]}(s_0)$ happens to be linearly dependent upon $F'(s_0)$, $F''(s_0)$, $\cdots$, $F^{[r]}(s_0)$, then

$$V_r'(s_0) = -k_{r-1}(s_0)V_{r-1}(s_0).$$

The coefficients appearing above, $k_1(s)$, $k_2(s)$, $\cdots$, $k_{r-1}(s)$, are the *curvatures* associated with the given curve at the point $F(s)$. The rth curvature, $k_r(s)$, may be defined similarly when $F^{[r+1]}(s)$ is independent of $F'(s)$, $F''(s)$, $\cdots$, $F^{[r]}(s)$, and to be zero in the dependent case. It turns out that $k_i(s)>0$ for $1\leqq i\leqq r-1$, and $k_r(s)\geqq 0$. For $1\leqq i\leqq r-1$, $k_i(s)$ will be of class C^{k-i-1}. $k_r(s)$ will be of class C^{k-r-1} wherever it does not vanish, but overall can only be guaranteed to be continuous. There is a standard way out of this dilemma in the special case $r=n-1$, for then there is a natural choice for $V_{r+1}(s)$. We do not go into details, but simply remark that the corresponding $k_r(s)$ will then be of class C^{k-r-1} on all of I, but will not necessarily be a positive function.

3. The algorithm for parametrizations by arc length. The Gram-Schmidt process is actually carried out as follows. Let

$$E_1(s) = F'(s) \qquad \text{and} \qquad V_1(s) = \frac{E_1(s)}{|E_1(s)|}.$$

If $V_1(s)$, $V_2(s)$, $\cdots$, $V_{i-1}(s)$ have already been determined, let

$$E_i(s) = F^{[i]}(s) - \textstyle\sum_{j<i}[F^{[i]}(s)\cdot V_j(s)]V_j(s)$$

and

$$V_i(s) = \frac{E_i(s)}{|E_i(s)|}.$$

This works for $i=2, 3, \cdots, r$. The process can actually be carried a half step further by computing

$$E_{r+1}(s) = F^{[r+1]}(s) - \textstyle\sum_{j<r+1}[F^{[r+1]}(s)\cdot V_j(s)]V_j(s),$$

since $r<k$. On the other hand, there is no guarantee that $E_{r+1}(s)\neq 0$, so in general we can not form $V_{r+1}(s)$. The vectors $E_1(s)$, $E_2(s)$, $\cdots$, $E_{r+1}(s)$ are conveniently referred to as the *excess vectors*, since $E_i(s)$ is nothing but the component of $F^{[i]}(s)$ orthogonal to the subspace spanned by $F'(s)$, $F''(s)$, $\cdots$, $F^{[i-1]}(s)$.

The algorithm for computing the curvatures derives from the following

THEOREM 3.1. *$k_i(s)=|E_{i+1}(s)|/|E_i(s)|$ for $1\leqq i\leqq r$.*

First assume $i<r$. Then

$$k_i(s) = V_i'(s)\cdot V_{i+1}(s) = \left(\frac{E_i(s)}{|E_i(s)|}\right)'\cdot V_{i+1}(s)$$

$$= \frac{E_i'(s) \cdot V_{i+1}(s)}{| E_i(s) |} + \left(\frac{1}{| E_i(s) |} \right)' E_i(s) \cdot V_{i+1}(s).$$

Now $E_i(s)$ and $V_{i+1}(s)$ are orthogonal, so the second term on the right above is zero. Hence

$$k_i(s) = \frac{E_i'(s) \cdot V_{i+1}(s)}{| E_i(s) |} .$$

To verify the theorem, we must show that $E_i'(s) \cdot V_{i+1}(s) = | E_{i+1}(s) |$.

Differentiating the equation $E_i(s) = F^{[i]}(s) - \sum_{j<i} [F^{[i]}(s) \cdot V_j(s)] V_j(s)$ yields

$$E_i'(s) = F^{[i+1]}(s) - \sum_{j<i}[F^{[i]}(s) \cdot V_j(s)]' V_j(s) - \sum_{j<i}[F^{[i]}(s) \cdot V_j(s)] V_j'(s).$$

Every vector on the right hand side of this last equation, except for $F^{[i+1]}(s)$, is a linear combination of $V_1(s), V_2(s), \cdots, V_i(s)$, and these are all orthogonal to $V_{i+1}(s)$. Therefore

$$E_i'(s) \cdot V_{i+1}(s) = F^{[i+1]}(s) \cdot V_{i+1}(s).$$

But $E_{i+1}(s) = F^{[i+1]}(s) - \sum_{j<i+1} [F^{[i+1]}(s) \cdot V_j(s)] V_j(s)$, so

$$F^{[i+1]}(s) \cdot V_{i+1}(s) = E_{i+1}(s) \cdot V_{i+1}(s) = | E_{i+1}(s) | ,$$

completing the proof in the case $i < r$.

If $i = r$ and $E_{i+1}(s) \neq 0$, the same proof works. If $E_{i+1}(s) = 0$, so is $k_i(s) = 0$, and the theorem is then certainly true.

4. The algorithm for arbitrary parametrizations. If $F^*: I^* \to R^n$ is a C^k-immersion (i.e., $F^{*\prime}(t) \neq 0$ for all $t \in I^*$) which is not necessarily a parametrization by arc length, the situation is only a trifle more involved. As before, we suppose that for each $t \in I^*$, the vectors

$$F^{*\prime}(t), F^{*\prime\prime}(t), \cdots, F^{*[r]}(t) \qquad r < k$$

are linearly independent. As in the preceding section, the Gram-Schmidt process yields the orthogonal excess vectors

$$E_1^*(t), E_2^*(t), \cdots, E_r^*(t), E_{r+1}^*(t)$$

and the orthonormal frame vectors

$$V_1^*(t), V_2^*(t), \cdots, V_r^*(t).$$

Let I be an interval in R^1 and $h: I \to I^*$ an orientation-preserving C^k-diffeomorphism such that $F = F^*h$ is a C^k-parametrization by arc length. Such a change of parameter always exists. Evaluating the various derivatives of F in terms of those of F^* via the chain rule, one easily sees that for $1 \leq i \leq r$,

$$F'(s), F''(s), \cdots, F^{[i]}(s) \quad \text{and} \quad F^{*\prime}(t(s)), F^{*\prime\prime}(t(s)), \cdots, F^{*[i]}(t(s))$$

generate the same subspace of R^n. In particular, $F'(s), F''(s), \cdots, F^{[r]}(s)$ are linearly independent for all $s \in I$, so we are in a position to calculate the orthogonal excess vectors

$$E_1(s), E_2(s), \cdots, E_r(s), E_{r+1}(s)$$

and the orthonormal frame vectors

$$V_1(s), V_2(s), \cdots, V_r(s)$$

for the parametrization F. Looking carefully at the chain rule formulas for the derivatives of F in terms of those of F^* leads one to the following conclusions.

THEOREM 4.1. *For* $1 \leqq i \leqq r+1$, $E_i(s) = E_i^*(t(s))(dt/ds)^i$ *and therefore* $|E_i(s)| = |E_i^*(t(s))| / |E_1^*(t(s))|^i$. *For* $1 \leqq i \leqq r$, $V_i(s) = V_i^*(t(s))$.

Let $k_i^*(t)$ denote the ith curvature of the curve defined by F^* at the point $F^*(t)$. Then from Theorems 3.1 and 4.1 we immediately get

THEOREM 4.2.

$$k_i^*(t) = \frac{|\overset{*}{E}_{i+1}(t)|}{|E_1^*(t)|\ |E_i^*(t)|} \qquad \textit{for} \quad 1 \leqq i \leqq r.$$

This gives the algorithm for computing the curvatures in the case of a parametrization not necessarily by arc length.

5. An example. Let p and q be real numbers, $0 < p < q$. Consider the curve given by $F: R^1 \to R^4$, where $F(t) = (\cos pt, \sin pt, \cos qt, \sin qt)$. This curve lies on the torus

$$T = \left\{ (x_1, x_2, x_3, x_4): \overset{2}{x_1} + \overset{2}{x_2} = 1, \overset{2}{x_3} + \overset{2}{x_4} = 1 \right\}$$

in four-space. The curve will be closed if and only if p/q is rational. Given any two points on the curve, there is an isometry of R^4 onto itself which takes the curve onto itself and takes the one point onto the other. Thus the various curvatures are the same at all points of the curves. They may be calculated according to the algorithm given by Theorem 4.2, as follows.

$$F'(0) = (0, p, 0, q) \qquad F'''(0) = (0, -p^3, 0, -q^3)$$
$$F''(0) = (-p^2, 0, -q^2, 0) \qquad F^{\mathrm{iv}}(0) = (p^4, 0, q^4, 0).$$

Beginning the Gram-Schmidt process, we get

$$E_1(0) = (0, p, 0, q)$$
$$|E_1(0)| = \sqrt{(p^2 + q^2)}$$
$$V_1(0) = \frac{1}{\sqrt{(p_2 + q^2)}}(0, p, 0, q).$$

Next stage:

$$E_2(0) = (-p^2, 0, -q^2, 0)$$
$$|E_2(0)| = \sqrt{(p^4+q^4)}$$
$$V_2(0) = \frac{1}{\sqrt{(p^4+q^4)}}(-p^2, 0, -q^2, 0).$$

Third Stage:

$$E_3(0) = \frac{pq(q^2-p^2)}{p^2+q^2}(0, q, 0, -p)$$
$$|E_3(0)| = \frac{pq(q^2-p^2)}{\sqrt{(p^2+q^2)}}$$
$$V_3(0) = \frac{1}{\sqrt{(p^2+q^2)}}(0, q, 0, -p).$$

Final stage:

$$E_4(0) = \frac{p^2q^2(q^2-p^2)}{p^4+q^4}(-q^2, 0, p^2, 0)$$
$$|E_4(0)| = \frac{p^2q^2(q^2-p^2)}{\sqrt{(p^4+q^4)}}$$
$$V_4(0) = \frac{1}{\sqrt{(p^4+q^4)}}(-q^2, 0, p^2, 0).$$

The curvatures are then given by:

$$k_1(0) = \frac{|E_2(0)|}{|E_1(0)|\,|E_1(0)|} = \frac{\sqrt{(p^4+q^4)}}{p^2+q^2} = \frac{\sqrt{(1+r^4)}}{1+r^2},$$

where $r=p/q$, and hence $0<r<1$.

$$k_2(0) = \frac{|E_3(0)|}{|E_1(0)|\,|E_2(0)|} = \frac{pq(q^2-p^2)}{(p^2+q^2)\sqrt{(p^4+q^4)}} = \frac{r(1-r^2)}{(1+r^2)\sqrt{(1+r^4)}},$$

$$k_3(0) = \frac{|E_4(0)|}{|E_1(0)|\,|E_3(0)|} = \frac{pq}{\sqrt{(p^4+q^4)}} = \frac{r}{\sqrt{(1+r^4)}}.$$

6. Miscellany. Several nice formulas follow easily from Theorems 3.1 and 4.2. We give them here only for parametrizations by arc length and leave the derivation of the corresponding formulas in the general case—actually no more difficult—to the reader.

Denote by $v_i(s)$ the hypervolume of the parallelepiped with "leading edges" $F'(s), F''(s), \cdots, F^{[i]}(s)$, so that

$$v_i(s) = |E_1(s)|\,|E_2(s)| \cdots |E_i(s)|.$$

Since $k_i(s)=|E_{i+1}(s)|/|E_i(s)|$, we have $|E_i(s)|=k_1(s)k_2(s)\cdots k_{i-1}(s)$. Hence

$$v_i(s) = [k_1(s)]^{i-1}[k_2(s)]^{i-2}\cdots[k_{i-2}(s)]^2[k_{i-1}(s)].$$

Also, since $|E_i(s)|=v_i(s)/v_{i-1}(s)$, we get

$$k_i(s) = \frac{v_{i+1}(s)v_{i-1}(s)}{[v_i(s)]^2}.$$

This last equation is essentially the same as formula (6.2-17) in J. C. H. Gerretsen, *Lectures on Tensor Calculus and Differential Geometry*, Noordhoff, Groningen, 1962.

The author holds an Alfred P. Sloan Research Fellowship.

THE KURATOWSKI CLOSURE PROBLEM IN THE TOPOLOGY OF CONVEXITY

WILLIAM KOENEN, Highland Park High School, St. Paul, Minnesota

Topologies are usually given in terms of open sets, but it is common also to speak of a topology associated with a closure operator. For example, see Kelley [6] page 43. Hammer [2, 3, 5] has extended the notions of topology and has defined an *extended topology* to be a pair (M, g), where M is some space and g is an expansive function defined on subsets of M. By *expansive*, we mean that for all A and B, subsets of M, $A\subseteq gA$ and $gA\cup gB\subseteq g(A\cup B)$. An expansive function g leads to a unique closure function [2] of a general type to be defined below. We could thus speak of an extended topology associated with a closure operator. In this manner, every situation in one extended topology suggests an analogue in another. Such is the case with Kuratowski's problem which is the subject of the present investigation. The usefulness of Hammer's functional notation is also illustrated. Whenever two closure operators act in the same space, as in this problem, the conventional notations are clumsy.

We shall consistently use M for the space and N for the null set. For $A\subseteq M$, the *complement operator* c is defined by the pair of equations, $cA\cap A=N$ and $cA\cup A=M$; cA is called the *complement of A in M*. Any operator k which satisfies the following four conditions [6] is called a Kuratowski closure operator.

THE KURATOWSKI CLOSURE AXIOMS: *For A, $B\subseteq M$:*

1. $kN=N$;
2. $A\subseteq kA$;
3. $k^2=k$, *where* $k^2A=k(kA)$;
4. $k(A\cup B)=kA\cup kB$.

For $S\subseteq M$, we define

$$\mathcal{C}_1(k, S) = \{S, kS, ckS, kckS, ckckS, kckckS, \cdots\},$$
$$\mathcal{C}_2(k, S) = \{cS, kcS, ckcS, kckcS, ckckcS, \cdots\},$$

and $\mathcal{C}(k, S)=\mathcal{C}_1(k, S)\cup\mathcal{C}_2(k, S)$. Observe that $\mathcal{C}_2(k, S)=\mathcal{C}_1(k, cS)$; hence the structure of $\mathcal{C}_2$ will be revealed in the study of $\mathcal{C}_1$.

It is possible to show that for a Kuratowski closure operator k, the infinity of symbols representing the elements of $\mathcal{C}_i(k, S)$ actually name only a finite number of distinct objects [4]. In fact, the cardinal number of $\mathcal{C}(k, S)$, which we denote by $\|\mathcal{C}(k, S)\|$, is at most fourteen, and by taking M to be the real line E^1 and k to be the usual closure operator of topology, one can find [1] $S\subseteq M$ for which $\|\mathcal{C}(k, S)\|=14$.

Hammer has shown that $\|\mathcal{C}(k, S)\|\leqq 14$ is a consequence of the weaker set of conditions [4] given below.

THE CLOSURE PROPERTIES: *For* $A, B\subseteq M$,

2. $A\subseteq kA$;
3. $k^2=k$;

4a. $kA\cup kB\subseteq k(A\cup B)$.

The condition 4a is sometimes stated in one of the following equivalent forms:

4b. $k(A\cap B)\subseteq kA\cap kB$;

4c. $A\subseteq B\Rightarrow kA\subseteq kB$.

Any operator which satisfies the weaker set of conditions is called simply a *closure operator* [2].

The *convex hull operator* h is defined by

$$hS = \bigcap \{X: S \subseteq X \subseteq M = E^n; \text{ and } X \text{ convex}\}.$$

The symbol hS is read *the convex hull of* S.

Henceforth we shall use h for the convex hull operator and k for the usual topological closure operator in $M=E^n$. The operator h is a closure, but not a Kuratowski closure operator, while k is both.

Since there are examples with $\|\mathcal{C}(k, S)\|=14$, it is natural to seek examples with $\|\mathcal{C}(h, S)\|=14$. Such a set S cannot be found.

THEOREM. *If* $M=E^n$ *and* $S\subseteq M$ *then* $\|\mathcal{C}(h, S)\|\leqq 10$; *moreover, there is a set* $S^*\subseteq E^2$ *for which* $\|\mathcal{C}(h, S^*)\|=10$.

The proof of the above theorem rests upon several preliminary results.

An *open half-space* in E^n is a set expressible in the form, $H=\{(x_1, x_2, x_3, \cdots, x_n): a_0+\sum_{i=1}^n a_ix_i<0; a_i \text{ real and not all } a_1, a_2, a_3, \cdots, a_n \text{ are zero}\}$. If H is an open half-space, then kH is a *closed half-space*, and any set D for which $H\subseteq D \subseteq kH$ is a *half-space.* Among the half-spaces, the so-called semi-spaces [5] are of special importance in the theory of convexity. A *semi-space* is a maximal convex

set which excludes a certain point. For example, if $p \in M = E^3$, let A be an open half-space whose face contains p; in the face of A, let B be an open half-plane whose edge contains p; and along the edge of B, let C be an open half-line whose origin is p. Then $A \cup B \cup C$ is a semi-space.

The *boundary operator* b is given by $bS = kS \cap kcS$ for $S \subseteq M$, and bS is called the *boundary of* S. The *interior* of S is defined by iS, where $i = ckc$.

The *direction* of a half-space is an outward normal unit vector. Half-spaces are *equally directed* if they have the same direction, *oppositely directed* if their directions, $\mathbf{d}_1$ and $\mathbf{d}_2$, satisfy $\mathbf{d}_1 + \mathbf{d}_2 = 0$.

We shall make use of the following results.

LEMMA 1. *If D is a half-space, then hD, kD, and iD are half-spaces with the same direction, while cD is a half-space with opposite direction.*

LEMMA 2. *If a union of equally directed half-spaces is not M, then it is a half-space. If an intersection of equally directed half-spaces is not N, then it is a half-space.*

LEMMA 3. *A maximal convex set excluding $p \in M$—i.e., a semi-space—is a half-space.*

THEOREM 1. *If D_1 and D_2 are two unequally directed half-spaces in $M = E^n$, then $h(D_1 \cup D_2) = M$.*

Proof. Let $\mathbf{d}_1 \neq \mathbf{d}_2$ be the directions of D_1 and D_2 respectively. If $D_1 \cup D_2 = M$, then $h(D_1 \cup D_2) = M$. On the other hand, if $D_1 \cup D_2 \neq M$, then there is a point $p \in cD_1 \cap cD_2$. Any ray from p in the direction $\mathbf{q} = \mathbf{d}_1 - \mathbf{d}_2 \neq 0$ will contain points of iD_2 because $\mathbf{q}$ has a positive component towards iD_2. Also, a ray from p in the direction $-\mathbf{q} = \mathbf{d}_2 - \mathbf{d}_1$ will contain points of iD_1. Since p lies between points of $D_1 \cup D_2$, $p \in h(D_1 \cup D_2)$. Therefore $h(D_1 \cup D_2) = M$.

THEOREM 2. *If D_1 and D_2 are two unequally directed half-spaces in $M = E^n$, and if $S \subseteq D_1 \cap D_2$, then $\|\mathfrak{C}_1(h, S)\| \leqq 5$.*

Proof. From closure properties, $hS \subseteq h(D_1 \cap D_2) \subseteq hD_1 \cap hD_2$.

Complementing, $chS \supseteq chD_1 \cup chD_2$. Since chD_1 and chD_2 are unequally directed half-spaces, and in view of Theorem 1, $hchS \supseteq h(chD_1 \cup chD_2) = M$. Therefore $hchS = M$; $chchS = N$; and $\mathfrak{C}_1(h, S) = \{S, hS, chS, M, N\}$. If all five element symbols name distinct objects, the cardinality is five. In any case $\|\mathfrak{C}_1(h, S)\| \leqq 5$.

THEOREM 3. *If S lies in no half-space of M, then $\|\mathfrak{C}_1(h, S)\| \leqq 3$.*

Proof. Since $S \subseteq hS$, hS is a convex set which lies in no half-space. Hammer has shown [5] that this guarantees that $hS = M$. Hence $\mathfrak{C}_1(h, S) = \{S, M, N\}$.

In view of Theorems 2 and 3, it is apparent that if $\|\mathfrak{C}_1(h, S)\|$ is to exceed five, S cannot lie in the intersection of two unequally directed half-spaces, nor can it fail to lie in at least one half-space. The possibility remaining is that S

lies in a half-space of direction **d**, but in no half-space of another direction. In this case, it lies in a family of half-spaces of direction **d**.

THEOREM 4. *If S lies in a half-space of direction* **d**, *but in no half-space of a different direction, then hS is a half-space.*

Proof. Hammer has shown that hS is the intersection of a family of semi-spaces [5]. Each of these semi-spaces contains S; hence each has direction **d**. Therefore hS is a half-space.

Observe that if S satisfies the conditions of Theorem 4, then hS, chS, $hchS$, $chchS$, etc., are all half-spaces, and whether or not they differ depends on those points in bhS. Notice also that bhS is E^{n-1}. The sets hS, chS, $hchS$, $chchS$, etc., in E^n differ exactly when the sets $hS\cap bhS$, $chS\cap bhS$, $hchS\cap bhS$, $chchS\cap bhS$, etc., in E^{n-1} differ.

We have shown that the cardinality of $\mathcal{C}_1(h, S)$ depends on the situation in a space of one less dimension if it is to exceed five. We shall use the reduction of dimension in an induction.

THEOREM 5. *If $S\subseteq M=E^1$, then $\|\mathcal{C}_1(h, S)\|\leqq 5$.*

Proof. In view of Theorems 2 and 3, we need only to consider the case where S is a subset of a positively directed half-line but not a subset of any half-line in the opposite direction. In this case, hS is a half-line and so is chS. Further application of the operators h and c merely reproduces the same two half-lines. Therefore $\|\mathcal{C}_1(h, S)\|$ in this case does not exceed three. In any case, the cardinality could not exceed five.

THEOREM 6. *For $S\subseteq M=E^n$, $\|\mathcal{C}_1(h, S)\backslash S\|\leqq 4$.*

Proof. As a direct application of Theorem 5, the assertion is true for $n=1$.

Suppose S lies in a half-space of direction **d**, but in no half-space of another direction, and suppose also that the assertion is true for $n-1$. Then all elements of $\mathcal{C}_1(h, S)\backslash S$ are half-spaces which are distinguishable only by the configuration of points in $bhS=E^{n-1}$. But by assumption, the cardinality for E^{n-1} cannot exceed four. The theorem follows by induction.

THEOREM 7. *For $S\subseteq M=E^n$, $\|\mathcal{C}_1(h, S)\|\leqq 5$.*

THEOREM 8. *For $S\subseteq M=E^n$, $\|\mathcal{C}(h, S)\|\leqq 10$.*

Proof. $\mathcal{C}(h, S)=\mathcal{C}_1(h, S)\cup\mathcal{C}_2(h, S)=\mathcal{C}_1(h, S)\cup\mathcal{C}_1(h, cS)$.

EXAMPLE. Let $M=E^2$, and let $S^*=\{(0, 1)\}\cup\{(x, y): y\leqq 0;\ x^2+y^2\neq 0\}$. Then

$$\begin{aligned} hS^* &= \{(0, 1)\} \cup \{(x, y): y < 1\}; \\ chS^* &= \{(x, y): y \geqq 1;\ x^2 + (y-1)^2 \neq 0\}; \\ hchS^* &= \{(x, y): y \geqq 1\}; \\ chchS^* &= \{(x, y): y < 1\}. \end{aligned}$$

Also

$$cS^* = \{(0, 0)\} \cup \{(x, y) : y > 0; x^2 + (y-1)^2 \neq 0\};$$
$$hcS^* = \{(0, 0)\} \cup \{(x, y) : y > 0\};$$
$$chcS^* = \{(x, y) : y \leqq 0; x^2 + y^2 \neq 0\};$$
$$hchcS^* = \{(x, y) : y \leqq 0\};$$
$$chchcS^* = \{(x, y) : y > 0\}.$$

The above sets are distinct.

The writer wishes to thank the referee for his suggestions, and Preston C. Hammer for his help in constructing the above example, and for suggesting the investigation into E^3 and E^n.

Work on this paper was supported by the NSF through the University of Wisconsin Research Participation Program, directed by Donald H. Bucklin.

References

1. Thomas A. Chapman, A further note on closure and interior operators, this MONTHLY, 69 (1962) 524–529.
2. P. C. Hammer, Extended topology: Set-valued set-functions, Nieuw Arch. Wisk., (3) 10 (1962) 55–57.
3. ———, General topology, symmetry, and convexity, Trans. Wisconsin Acad. Sci. Arts Lett., 44 (1955) 221–225.
4. ———, Kuratowski's closure theorem, Nieuw Arch. Wisk., 8 (1960) 74–80.
5. ———, Semispaces and the topology of convexity, Amer. Math. Soc., Proc. Symposia in Pure Math., 7 (1963) 305–316.
6. J. L. Kelley, General topology, Van Nostrand, New York, 1955.
7. O. Ore, Theory of graphs. Amer. Math. Soc. Colloq. Pub., 38 (1962) (see esp. p. 177).

POLYNOMIAL REMAINDER SEQUENCES AND DETERMINANTS

G. E. COLLINS, IBM Watson Research Center, Yorktown Heights, N. Y.

Introduction. Let P_1 and P_2 be polynomials over an arbitrary Gaussian domain $\mathscr{I}$ with $\deg(P_1) \geqq \deg(P_2) > 0$, where $\deg(P)$ denotes the degree of P. We consider the sequence of polynomials $P_1, P_2, P_3, \cdots$ in which P_{i+2} is the remainder obtained when P_i is divided by P_{i+1} by means of the Euclidean algorithm. We will show that, to within a constant factor, the coefficients of the P_i, $i > 2$, can be expressed as certain determinants in the coefficients of P_1 and P_2. A formula is obtained which relates the polynomial remainders defined by the determinants to those obtained by means of the Euclidean algorithm. This formula reveals that, in the typical case, the Euclidean algorithm systematically introduces extraneous constant factors. The computational problem of eliminating such factors, in case $\mathscr{I}$ is the domain I of the integers or a polynomial domain $I[x_1, \cdots, x_n]$ over I, is briefly discussed.

Definitions. We adopt the standard terminology that polynomials P and Q over $\mathfrak{s}$ are *associates* in case $aP=bQ$ for some $a, b\in\mathfrak{s}$ with $a\neq 0$, $b\neq 0$. Also, P is *primitive* in case any common divisor of its coefficients is a unit in $\mathfrak{s}$. Deg(P) will denote the degree of P.

Let $P(x)=\sum_{i=0}^{m} a_ix^i$ and $Q(x)=\sum_{i=0}^{n} b_ix^i$ be polynomials over $\mathfrak{s}$ of degrees m and n. If $m\geqq n$ and $Q\neq 0$, define $\rho(P, Q)=R$ where $R(x)=b_nP(x)-a_mx^{n-m}Q(x)$. If $m<n$ or if $Q=0$, define $\rho(P, Q)=P$. Define $\rho^0(P, Q)=P$ and inductively $\rho^{i+1}(P, Q)=\rho(\rho^i(P, Q), Q)$. The *Euclidean algorithm* for the remainder $\mathfrak{R}(P, Q)$ obtained by dividing P by Q is expressed by the formula $\mathfrak{R}(P, Q)=\rho^k(P, Q)$, where k is the least nonnegative integer such that $\rho^{k+1}(P, Q)=\rho^k(P, Q)$. The integer k so defined will be called the *rank* of P over Q and will be denoted by $r(P, Q)$.

Clearly $\mathfrak{R}(P, Q)=P$ if $Q=0$ or if $\deg(P)<\deg(Q)$, and $\mathfrak{R}(P, Q)=0$ if $\deg(Q)=0$ and $Q\neq 0$. Thus the only case of interest is $\deg(P)\geqq\deg(Q)>0$. In this case $aP=QS+R$ for some nonzero $a\in\mathfrak{s}$ ($a=b^k$, where b is the leading coefficient of Q and $k=r(P, Q)$) and some polynomial S over $\mathfrak{s}$, where $R=\mathfrak{R}(P, Q)$ and $\deg(R)<\deg(Q)$. We generalize this situation and say that (P, Q, R) is a *polynomial remainder triple* in case $\deg(P)\geqq\deg(Q)>\deg(R)$ and $aP=QS+bR$ for some $a, b\in\mathfrak{s}$, $a\neq 0$, $b\neq 0$, and some polynomial S over $\mathfrak{s}$. It is easy to verify that if (P, Q, R) is a polynomial remainder triple and P', Q' and R' are respective associates of P, Q and R, then (P', Q', R') is a polynomial remainder triple. Also, if (P, Q, R) and (P', Q', R') are polynomial remainder triples and P' and Q' are respective associates of P and Q, then R' is an associate of R.

A polynomial remainder triple (P, Q, R) such that $R=\mathfrak{R}(P, Q)$ will be called *Euclidean*. A sequence of polynomials $P_1, P_2, \cdots, P_k$ $(k\geqq 3)$ such that (P_i, P_{i+1}, P_{i+2}) is a (Euclidean) polynomial remainder triple for $1\leqq i\leqq k-2$ will be called a *(Euclidean) polynomial remainder sequence*.

For the convenient statement and proof of our theorem we need also the following definition. Let M be a matrix over $\mathfrak{s}$ with m rows and n columns, $m\leqq n$. We associate with M a polynomial $P=\mathcal{P}(M)$, $\deg(P)\leqq n-m$, as follows.

For $0\leqq i\leqq n-m$ let M_i be the m by m square submatrix of M consisting of the first $m-1$ columns of M followed by the $(m+i)$-th column of M. Let a_i be the determinant of M_i. Then $P(x)=\sum_{i=0}^{n-m} a_ix^{n-m-i}$.

Let $P_1, P_2, \cdots, P_k$ be a Euclidean polynomial remainder sequence. Let $n_i=\deg(P_i)$ for $1\leqq i\leqq k$. Then $n_1\geqq n_2>n_3>\cdots>n_k\geqq 0$. We define the following matrix M, where $P_1(x)=\sum_{j=0}^{m} a_jx^j$, $P_2(x)=\sum_{j=0}^{n} b_jx^j$, $m=n_1$, and $n=n_2$.

$$M=\begin{bmatrix} a_m & a_{m-1} & a_{m-2} & \cdots & a_0 & 0 & 0 & \cdots & 0 \\ 0 & a_m & a_{m-1} & \cdots & a_1 & a_0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0\cdots & 0 & a_m & a_{m-1} & \cdots & \cdots & \cdots & \cdots & a_0 \\ b_n & b_{n-1} & b_{n-2} & \cdots & b_0 & 0 & 0 & \cdots & 0 \\ 0 & b_n & b_{n-1} & \cdots & b_1 & b_0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0\cdots & 0 & b_n & b_{n-1} & \cdots & \cdots & \cdots & \cdots & b_0 \end{bmatrix}$$

The matrix M consists of $n-n_{k-1}+1$ rows of P_1 coefficients a_j and $m-n_{k-1}+1$ rows of P_2 coefficients b_j. The number of columns of M is $m+n-n_{k-1}+1$. We will refer to M as the *associated matrix* of $P_1, P_2, \cdots, P_k$.

The Main Theorem.

THEOREM 1. *Let $P_1, P_2, \cdots, P_k$ be a Euclidean polynomial remainder sequence and let M be its associated matrix. Then P_k and $\mathcal{P}(M)$ are associates. More specifically, let $n_i=\deg(P_i)$ and let c_i be the leading coefficient of P_i, for $1\leqq i<k$. Let $r_i=r(P_{i-2}, P_{i-1})$ for $3\leqq i\leqq k$. Let $p=n_{k-1}-1$. Then*

$$\left(\prod_{i=2}^{k-2} c_i^{n_{i-1}-n_i+1}\right)\cdot c_{k-1}^{n_{k-2}-p}\cdot P_k = \pm\left(\prod_{i=2}^{k-1} c_i^{r_{i+1}(n_i-p)}\right)\cdot\mathcal{P}(M).$$

Proof. The essence of the proof is that all the computations required to compute P_k can be translated into successive elementary row operations performed on the matrix M. And any elementary row operation performed on M results in a new matrix M' such that $\mathcal{P}(M')$ is a constant multiple of $\mathcal{P}(M)$. Ultimately one obtains from M a matrix M^* one of whose rows consists of the coefficients of P_k. Evaluation of $\mathcal{P}(M^*)$ leads to the above formula.

We begin by multiplying each row of P_1 coefficients a_j by b_n. The result is a matrix M_1 such that

$$\mathcal{P}(M_1) = b_n^{n-p}\cdot\mathcal{P}(M) = c_2^{n_2-p}\cdot\mathcal{P}(M),$$

since there were $n-p=n_2-p$ rows of P_1 coefficients. Next we subtract a_m times the ith row of P_2 coefficients from the ith row of coefficients in M_1, for $1\leqq i\leqq n-p$. The result is a matrix M_2 which consists of $n-p$ rows of coefficients of $\rho(P_1, P_2)$ followed by $m-p$ rows of coefficients of P_2. Also, $\mathcal{P}(M_2)=\mathcal{P}(M_1)$. We note that the leading coefficient of $\rho(P_1, P_2)$ in the first row of M_2 occurs in the column $1+(m-\deg(\rho(P_1, P_2)))$, with leading coefficients of $\rho(P_1, P_2)$ in successive rows occurring in successive columns.

If $r_3=r(P_1, P_2)>1$, then a similar transformation is applied to M_2, producing a matrix M_4 whose first $n-p$ rows contain coefficients of $\rho^2(P_1, P_2)$. This time however the rows of coefficients of P_2 multiples of which are subtracted are the $(1+(m-\deg(\rho(P_1, P_2))))$-th row and the following $n-p-1$ rows of P_2 coefficients. The result is a matrix M_4 such that

$$\mathcal{P}(M_4) = c_2^{n-p}\cdot\mathcal{P}(M_2) = c_2^{2(n-p)}\cdot\mathcal{P}(M).$$

Repeating this process r_3 times one finally obtains a matrix $M_{2r_3}=M^{(3)}$, consisting of $n-p$ rows of coefficients of P_3 followed by $m-p$ rows of coefficients of P_2, such that $\mathcal{P}(M^{(3)})=c_2^{r_3(n-p)}\cdot\mathcal{P}(M)=c_2^{r_3(n_2-p)}\cdot\mathcal{P}(M)$.

If $k>3$, then the process which produced $M^{(3)}$ from M is repeated, with slight changes, to obtain a matrix $M^{(4)}$ containing coefficients of P_4. Specifically $M^{(4)}$ will consist of $n-p$ rows of P_3, followed by n_1-n_3 rows of P_2, followed by n_3-p rows of P_4, the first n_1-n_3 rows of P_2 in $M^{(3)}$ being left intact. And we have

$$\mathcal{P}(M^{(4)}) = c_3^{r_4(n_3-p)} \cdot \mathcal{P}(M^{(3)}).$$

If $k>4$, then the last n_4-p rows of P_3 coefficients are transformed into rows of P_5 coefficients by a similar process, using some of the rows of P_4 coefficients. The result is a matrix $M^{(5)}$ consisting of n_2-n_4 rows of P_3 coefficients, n_4-p rows of P_5 coefficients, n_1-n_3 rows of P_2 coefficients, and n_3-p rows of P_4 coefficients, in this order scanning from top to bottom.

Again, $\mathcal{P}(M^{(5)}) = c_4^{r_5(n_4-p)} \cdot \mathcal{P}(M^{(4)})$.

In general, as one may verify by induction on i, $M^{(2i)}$ will consist of n_2-n_4 rows of P_3, n_4-n_6 rows of P_5, $\cdots$, $n_{2i-4}-n_{2i-2}$ rows of P_{2i-3}, $n_{2i-2}-p$ rows of P_{2i-1}, n_1-n_3 rows of P_2, n_3-n_5 rows of P_4, $\cdots$, $n_{2i-3}-n_{2i-1}$ rows of P_{2i-2}, $n_{2i-1}-p$ rows of P_{2i} in this order, for $2i \leqq k$, while $M^{(2i+1)}$ will consist of n_2-n_4 rows of P_5, n_4-n_6 rows of P_5, $\cdots$, $n_{2i-2}-n_{2i}$ rows of P_{2i-1}, n_{2i-p} rows of P_{2i+1}, n_1-n_3 rows of P_2, n_3-n_5 rows of P_4, $\cdots$, $n_{2i-3}-n_{2i-1}$ rows of P_{2i-2}, $n_{2i-2}-p$ rows of P_{2i} in this order, for $2i+1 \leqq k$. For $j \leqq k$ we have $\mathcal{P}(M^{(j)}) = c_{j-1}^{r_j(n_{j-1}-p)} \cdot \mathcal{P}(M^{(j-1)})$. Combining all these equations we obtain

$$\mathcal{P}(M') = \left(\prod_{i=2}^{k-1} c_i^{r_{i+1}(n_i-p)} \right) \cdot \mathcal{P}(M). \tag{1}$$

By a suitable rearrangement of the rows of M' we obtain a matrix M'' such that the first nonzero element of each row, which is the leading coefficient c_i of some polynomial P_i is on the diagonal of M''. Hence by the definition of $\mathcal{P}$, we have

$$\mathcal{P}(M'') = \left(\prod_{i=2}^{k-2} c_i^{n_{i-1}-n_{i-3}} \right) \cdot c_{k-1}^{n_{k-2}-p} \cdot P_k. \tag{2}$$

Combining (1) and (2) together with $\mathcal{P}(M'') = \pm \mathcal{P}(M')$ produces the conclusion of the theorem.

A special case. One may suppose that ordinarily $\deg(\rho(P, Q)) = \deg(P)-1$, if $\deg(P) \geqq \deg(Q)$ and $Q \neq 0$; hence that $r(P, Q) = \deg(P) - \deg(\mathcal{R}(P, Q))$. Accordingly let us call a polynomial remainder sequence $P_1, P_2, \cdots, P_k$ *regular* in case $r_{i+2} = n_i - n_{i+2}$ for $1 \leqq i \leqq k-2$, where the n_i and r_i are defined as in Theorem 1. For a regular sequence clearly $n_{i+1}-n_{i+2} = 1$ for $1 \leqq i \leqq k-2$. Hence we have $n_{i-1}-n_{i+1} = 2$ for $3 \leqq i \leqq k-2$, $n_1-n_3 = n_1-n_2+1$, $n_i-p = k-i$ for $2 \leqq i \leqq k-1$, $r_i = 2$ for $i \leqq 4 \leqq k$, and $r_3 = n_1-n_2+1$. Making all these substitutions in the formula of Theorem 1 we obtain, after simplification:

COROLLARY 1. *Let $P_1, P_2, \cdots, P_k$ be a regular Euclidean polynomial remainder sequence. Then, in the notation of Theorem* 1,

$$P_k = \pm\, c_2^{(n_1-n_2+1)(k-3)} \left(\prod_{i=3}^{k-2} c_i^{2(k-i-1)} \right) \cdot \mathcal{P}(M).$$

In particular we see that in the regular case, with $k \geqq 4$, $\mathcal{P}(M)$ is a proper divisor of P_k (unless c_2 is a unit), in the sense that $P_k = a \cdot \mathcal{P}(M)$ where a is not

a unit. However, in the nonregular case it is easy to obtain a counter example to this assertion, even with $k=4$. In fact, one can construct an example for which P_4 is a proper divisor of $\mathcal{P}(M)$.

REMARKS. Our study has been motivated by the problem of computing polynomial remainder sequences in the case that $\mathcal{I}$ is $I[x_1, \cdots, x_n]$ (or, with $n=0$, I), I being the integral domain of the integers. Such a problem arises, for example, in the application of Tarski's decision method for the elementary theory of real-closed fields [3]. The problem also arises in computing the greatest common divisor of two polynomials. In this connection the reader may be interested in [4]. One would like to compute a polynomial remainder sequence $P_1, P_2, \cdots, P_k$ such that each P_i is primitive, and at the same time reduce as far as possible the amount of computing required.

The computing of the polynomials $\mathcal{P}(M)$ from the definition of the determinant function would be impractically laborious. On the other hand the computing of these determinants by the method of elimination amounts essentially to using the Euclidean algorithm followed by application of the formula of Theorem 1, a method which also requires too much computation.

The best algorithm which seems to be suggested by Theorem 1 is the following. Compute P_3 and P_4 by the Euclidean algorithm. Then divide P_4 as often as possible by c_2, and denote the result by P_4^*. Compute $P_5 = \mathcal{R}(P_3, P_4^*)$ by the Euclidean algorithm. Divide P_5 as often as possible by each of c_2 and c_3, and denote the result by P_5^*. Continue in this manner.

Another possibility would be to ignore Theorem 1 almost completely and, in the algorithm above, compute P_i^* as a primitive associate of P_i by computing the greatest common divisor of the coefficients of P_i. However, the only presently known practical method for computing such a greatest common divisor itself involves polynomial remainder sequences if the coefficients are polynomials, and the method soon becomes too laborious.

The special case of Corollary 1 obtained when one takes k such that $\mathcal{R}(P_{k-1}, P_k) = 0$, and takes $\mathcal{I}$ to be the field of real or complex numbers, has appeared previously in the literature, with the restriction that the factor relating P_k and $\mathcal{P}(M)$ has not been explicitly given. See, for example, [2], pp. 134–139 and [1], pp. 197–198.

Acknowledgment. I should like to express my indebtedness to Professor A. Robinson for some remarks which led to the present investigation.

References

1. M. Bôcher, Introduction to Higher Algebra, Macmillan, New York, 1907.

2. M. Marden, The geometry of the zeros of a polynomial in a complex variable, Amer. Math. Soc., Math. Survey No. 3, 1949.

3. A. Tarski, A decision method for elementary algebra and geometry, 2nd ed., Univ. of California Press, 1951.

4. W. S. Brown, J. P. Hyde and B. A. Tague, The ALPAK system for nonnumerical algebra on a digital computer, II, Rational functions of several variables and truncated power series with rational-function coefficients, The Bell System Technical J., March 1964, 785–804.

SOME BOMBING PROBLEMS

D. C. GILLILAND, Michigan State University

1. Introduction. Suppose that in Euclidean n-space a target T (assumed measurable) is to be attacked with a point bomb. Assume further that the bomb impact density is spherical normal, variance unity, about the aim point A. The problem is to aim the bomb so as to maximize the probability $\nu_A(T)$ of an impact within the target (a hit). In [1] it is shown that if T consists of two nonoverlapping hyperspheres of equal radii centered within two units of each other then $\nu_A(T)$ is maximum if and only if A coincides with the midpoint of the line segment between the centers. In the following this result will be referred to as (R). This paper reports extensions of (R) to more general symmetric targets. The following problem is also considered. If the center of the spherical normal density is fixed, how should k hyperspheres of equal radii be placed so as not to overlap and so as to maximize the probability measure of their union? For 2-space this has been referred to as a cookie cutter problem. Given a sheet of dough with thickness proportional to the circular normal density, how should one punch out k circular cookies so as to get the greatest amount of dough? If the no-overlap constraint is removed the problem is more difficult. In this case graphs for determining the optimal placement of two disks (spheres) have been published by Marsaglia [2].

In what follows O denotes the origin, P a generic point, and $|PO|$ the Euclidean distance between P and O. The probability of hitting a hypersphere of radius r centered a distance z from the aim point is given by

$$g(z) = z^{1-(1/2)n} \exp\{-\tfrac{1}{2}z^2\} \int_0^r t^{(1/2)n} \exp\{-\tfrac{1}{2}t^2\} I_{(1/2)n-1}(zt)\,dt,$$

where I_η denotes the modified Bessel function of the first kind and order η [4]. Throughout this paper the term hypersphere is to mean nondegenerate hypersphere.

2. Target symmetric about a point.

PROPOSITION 1. *If a target T consisting of nonoverlapping hyperspheres centered in the region $|PO| \leqq 1$ is symmetric about O, then $\nu_A(T)$ is maximum if and only if $A = O$.*

Proof. The hyperspheres not centered at O must occur in symmetric pairs. By (R) the contribution to $\nu_A(T)$ from each such pair is maximum if and only if $A = O$.

PROPOSITION 2. *If a target T contained in the region $|PO| \leqq 1$ is symmetric about O, then $\nu_A(T)$ is maximum if $A = O$.*

Proof. Let $\delta > 0$ be given. There exists an open set $H \supset T$ such that $\mu(H-T) < \delta$ where μ denotes Lebesgue measure. Consider H^* the intersection of H and $|PO| < 1$. Apply the Vitali covering theorem [3, p. 69] with the cover $\mathcal{S}$ consist-

ing of all closed hyperspheres contained in H^* to obtain a set $S=\bigcup_i S_i$ (a disjoint union with $\{S_i\}\subset\mathcal{S}$) such that $\mu(H^*-S)=0$. Since $\nu_A(B)\leqq\mu(B)$ for any measurable set B and any point A,

$$0 \leqq \nu_A(S) - \nu_A(T) \leqq \nu_A(H) - \nu_A(T) \leqq \mu(H - T) < \delta.$$

S can be constructed to be symmetric about O by working with one orthant at a time and reflecting through O. By Proposition 1 $\nu_A(S)$ is maximum if and only if $A=O$. Hence $\nu_A(T)$ is maximized when $A=O$.

3. Target symmetric about a line.

PROPOSITION 3. *Let T consist of nonoverlapping hyperspheres centered at the points P_i and let C denote the convex hull of the P_i. Suppose T is symmetric about a line L. If $|PP_j|+|PP_{j'}|\leqq 2$ for all $P\in C$ and P_j, $P_{j'}$ the centers of hyperspheres symmetric about L, then the aim point A that maximizes $\nu_A(T)$ must lie on L.*

Proof. (The author is indebted to the referee, whose remarks led to the formulation of this proposition.) The monotone property of g indicates that $\nu_A(T)$ is maximized for some $A\in\overline{C}$, the closure of C. Since T is symmetric about L those hyperspheres not centered on L must occur in pairs symmetric about L. For P_1 and P_2 the centers of any such pair and $A\in\overline{C}-L$ let $|AP_1|+|AP_2|=a$. On the ellipsoid $|PP_1|+|PP_2|=a$ the contribution to $\nu_P(T)$ from the hyperspheres at P_1 and P_2 is maximum if and only if $|PP_1|=|PP_2|$ since $a\leqq 2$. This follows from the argument given in [1] for P constrained to the line segment (degenerate ellipsoid) between P_1 and P_2. Thus given any point $A\in\overline{C}-L$ the projection A' of A on L will yield greater probability of a hit, i.e., $\nu_{A'}(T)>\nu_A(T)$.

We state a corollary which will be needed in the next section.

COROLLARY 1. *If T consists of three nonoverlapping hyperspheres of equal radii centered at the vertices of an equilateral triangle of perimeter no greater than three, then $\nu_A(T)$ is maximum if and only if A coincides with the center of the triangle.*

Proof. T is symmetric about the medians so $\nu_A(T)$ is maximum if and only if A coincides with their intersection.

We note that Proposition 3 obtains with L a subspace and that this strengthened version can be extended to measurable targets.

4. Hyperspheres free to move. Now let us suppose that k hyperspheres of radii r are free to move, the center of the probability distribution is fixed, say at O, and the arrangement of nonoverlapping hyperspheres that maximizes the probability measure of their union is sought.

PROPOSITION 4. *If $r\leqq 1$ the probability measure of the union of two nonoverlapping hyperspheres is maximum if the hyperspheres are tangent at the point O.*

Proof. The proof follows directly from (R).

PROPOSITION 5. *If $r \leqq 1/2$, $n > 1$, the probability measure of the union of three nonoverlapping hyperspheres is maximum if the hyperspheres are located at the vertices of an equilateral triangle of perimeter* $6r$ *centered at O.*

Proof. Clearly an optimal arrangement will have each tangent to the other two. The proof is completed by an appeal to Corollary 1.

PROPOSITION 6. *In 2-space if r is sufficiently small the probability measure of the union of four nonoverlapping disks is maximum if the disks are centered at the corners of a rhombus with diagonal lengths* $(2r, 2\sqrt{3}r)$ *and the center of the rhombus is at O.*

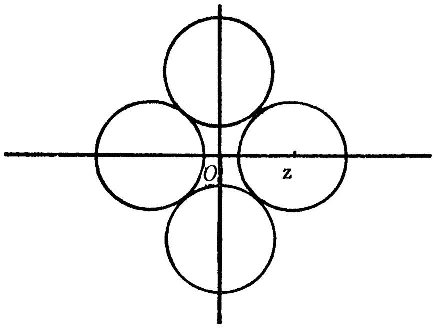


FIG. 1

Proof. An optimal arrangement results when the disk centers are at the corners of a rhombus with side length $2r$. By Proposition 1 the center of the rhombus must be at O for sufficiently small r. From Figure 1 we conclude that the probability measure of the disks is given by

$$f(z) = 2g(z) + 2g\sqrt{(4r^2 - z^2)}, \qquad r \leqq z \leqq \sqrt{3}r,$$

where g is as previously defined with $n = 2$. Thus

$$f'(z) = 2g'(z) - \frac{2z}{\sqrt{(4r^2 - z^2)}} g'\sqrt{(4r^2 - z^2)}$$

and from [1] $g'(z) = -r \exp\{-\frac{1}{2}(r^2+z^2)\} I_1(rz)$. Using the series expansion of I_1 [5] we can write

$$f'(z) = zr^2 \exp\{-\tfrac{1}{2}(r^2 + z^2)\} h(z), \tag{1}$$

where

$$h(z) = \sum_{j=0}^{\infty} \frac{r^{2j}}{4^j j!(j+1)!} [(4r^2 - z^2)^j \exp\{z^2 - 2r^2\} - z^{2j}].$$

Now

$$(2)\quad h'(z) = 2z\exp\{z^2 - 2r^2\} + \sum_{j=1}^{\infty} \frac{2zr^{2j}}{4^{j}j!(j+1)!}\left[(4r^2 - z^2)^{j-1}(4r^2 - z^2 - j)\exp\{z^2 - 2r^2\} - jz^{2j-2}\right].$$

It is easy to show that the series in (2) is bounded in absolute value by $3r^3 \exp\{r^4\}$ for $0 < r \leqq z \leqq r\sqrt{3}$ and sufficiently small r. Therefore, it follows that

$$(3)\qquad h'(z) \geqq 2r\exp\{-r^2\} - 3r^3\exp\{r^4\};$$

and hence $h'(z) > 0$ for $0 < r \leqq z \leqq r\sqrt{3}$ and sufficiently small r. But $h(r\sqrt{2}) = 0$ so $h(z) < 0$ for $r \leqq z < r\sqrt{2}$, $h(z) > 0$ for $r\sqrt{2} < z \leqq r\sqrt{3}$. Since the sign of $f'(z)$ is the sign of $h(z)$ we see that $f(z)$ is maximized at the end points of the interval $r \leqq z \leqq r\sqrt{3}$, and the proposition is proved.

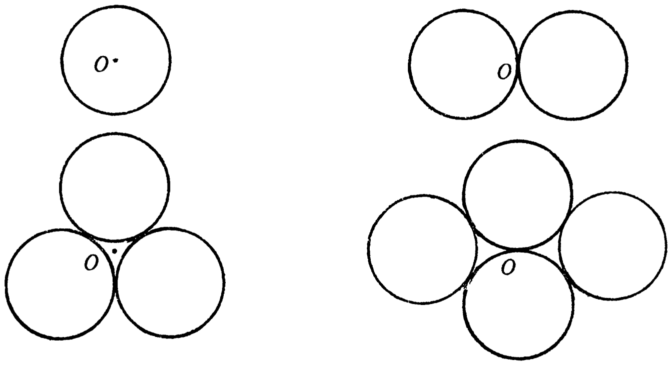


FIG. 2

These results give a solution of the cookie cutter problem for $k = 1, 2, 3$, or 4 circular cookies and small r. Figure 2 depicts optimal arrangements for each of these cases.

Many of the results reported in this paper depend upon the underlying probability measure only through (R). Loosely speaking, similar results hold for any measure for which (R) or a similar property obtains.

This work was supported in part by the National Science Foundation, Grant No. GP-6035.

References

1. D. C. Gilliland, A note on the maximization of a non-central chi-square probability, Ann. Math. Statist., 35 (1964) 441–442.

2. G. Marsaglia, Some problems involving circular and spherical targets, Operations Res., 13 (1965) 18–27.

3. W. W. Rogosinski, Volume and integral, Oliver and Boyd, London, 1952.

4. H. Ruben, Probability content of regions under spherical normal distributions, I, Ann. Math. Statist., 31 (1960) 598–618.

5. G. N. Watson, A treatise on the theory of Bessel functions, Cambridge University Press, 1922.

A NOTE ON THE GAME OF DOTS

J. C. HOLLADAY, Institute for Defense Analyses

The game of dots along with tick-tack-toe is a part of the lore of childhood although it may also be of interest to adults. Tick-tack-toe may be analyzed within a few hours and so, for many adults, it does not offer a sufficient challenge to maintain interest. The skills commonly developed for the game of dots are easily attained. Unlike tick-tack-toe, however, this game is known not to be trivial. The purpose of this paper is to present some results which are important for a better understanding of the game of dots.

The game of dots is played with a rectangular array of dots that form the vertices of congruent square cells called *squares*. Thus, if there are mn dots arranged in m rows and n columns, there will be $(m-1)(n-1)$ squares. The game is played by two players alternately connecting adjacent pairs of dots. When three sides of a square are filled in, however, the player whose turn it is must fill in the fourth side of the square, and then make another play. When all the squares are completed, the game is finished and the player who completes the greatest number of squares is the winner.

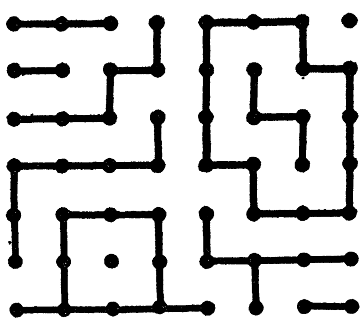


FIG. 1.—A typical position for which there are no "free" connections.

As the game is generally played, each player follows a strategy in which he tries to make connections, called *free* connections, that do not permit his opponent to complete squares. A player who is inept at finding such plays may find himself unnecessarily giving his opponent opportunities to score. If two players follow the above strategy, they will eventually arrive at a position (of which that of Figure 1 is typical) where there are no more "free" connections available. A player who is unskilled with this elementary strategy is generally one who believes that this stage of the game is reached before it has in fact been reached. When this second stage of the game arises, the players try to recognize the connections that offer the fewest squares for completion.

The unfilled squares in positions allowing no "free" connections are seen to be arranged in a way to suggest corridors. It is of interest to discuss these corridors in a manner that ignores both their lengths and their bends, in other words, topologically. To do this, let the center of each unplayed connection and the

center of each adjacent square be connected with a line interval. The lines formed by this process applied to Figure 1 are shown in Figure 2. The lines of Figure 1 are reproduced in Figure 2 as dotted lines.

The pattern of Figure 2 falls into six connected parts, called components. Two of the components are, topologically, circles, two are line intervals, and the other two are trees; i.e., connected line figures that, although having branch points, have no subsets that are circles. One of the trees has a three-branch point, and the other has both a three-branch point and a four-branch point.

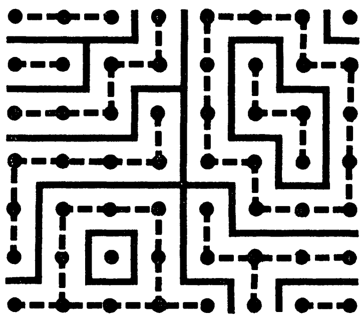


FIG. 2.—The topological nature of Figure 1 illustrated with solid lines.

Consider the not uncommon situation where the components of a position are all either circles, lines or trees. Then if a player on a given play sets up a collection of squares for his opponent to complete on his next move, the collection of squares will consist of either a line of squares or a circle of squares. Each end of such a line of squares is either unattached or leads to a branch square.

Any remaining collection of squares that may be immediately set up after a collection is filled in is the same as before but for the following exception: Whenever a line of squares has a three-branch square at the end of the line, the branch square ceases to be such a branch square and its two branches are thereby connected into a single line. Therefore, if a player, required to set up a collection of squares to be completed, sets up the smallest collection possible, any remaining collection to be set up will be at least as large. Let this situation be referred to as the *monotonicity result.*

If in such a position a player is assured of getting the last turn, he may obtain at least a draw (and more likely a win) if at each turn he sets up the least possible number of squares for his opponent. Which player goes last depends on which player goes first and whether there is an odd or even number of turns to the game. Let P be the number of plays (i.e., connections) to be made, and let S be the number of squares. Then the total number of turns is

$P-$number of replays in turns
$=P-$(number of plays that complete at least one square-1)
$=P-S+1+$number of plays that complete two squares
$=P-S+1+$the number of circles in the topological pattern.

Therefore, which player gets the last turn depends entirely upon whether the number of circles is even or odd.

If a position does not consist topologically only of circles, lines or trees, the analysis may become more complicated. Whenever a connection is made that sets up squares to be completed, the connection has two possible directions of influence. The completion of squares that occur in the next turn as a result of this play may be made by first progressing in one direction from this connection until no more squares may be thus completed, and then progressing in the other direction. The progression of completing squares may be described inductively as follows:

A direction from a newly made connection will satisfy exactly one of the following four cases:

A. *There is an adjacent square there and it has exactly one remaining connection to be made on it.*

B. *There is a square there with more than one remaining connection to be made.*

C. *There is no square there.*

D. *There is a square there with no remaining connection, so the square has just been completed.*

When Case A arises, the square is completed next and the other direction of the connection thus formed is studied in turn, and so the inductive process continues. When either Case B or Case C arises, no more completions are possible from this progression. If the other direction has not yet been studied, it is then likewise followed in this inductive manner.

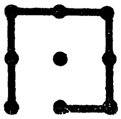


FIG. 3.—Four squares that, although capable of forming a circle, may be filled in as a line.

If at any prior stage of the game an opportunity was presented for completion of a square, it had to be completed at that time. Therefore, if Case D arises, it can only mean that one of the other sides of the square was completed by the original connection. In this case, no more squares are available for completion and a circle of squares has been completed. For the other cases, where the terminations of completions are accomplished by Case B or C or both, the process may be referred to as making a line of completions. As an example of such considerations, if the lower left connection of Figure 3 is made, a circle is formed. However, if some other connection is made, the four squares will be immediately filled in as a line. Such a "sacrificial" play might be made by a player wishing to maintain an even or odd number of circles.

When a general topological pattern is considered, the monotonicity result is not always valid. For instance, consider a component of squares topologically

like the letter P. Let the tail of the P represent more squares than the loop and branch square together. If the squares of the loop are set up for completion at this stage, then all of the squares of the whole P will be completed at once as a line. Therefore, in order to set up the least number of squares at this stage, the tail must be set up. After the squares of the tail are completed, the squares of the remaining circle contain fewer squares than the tail did.

Circles, lines and trees are not the only topological configurations for which the monotonicity result is valid. For each of the topological patterns of Figure 4, the monotonicity result is valid. Furthermore, for each of these configurations, no matter in what order their squares are completed, one and only one circle of squares will be completed. The properties of these patterns stem from the following characteristics which they have in common: both ends of each line lead to branch points, and no pattern contains two disjoint circles.

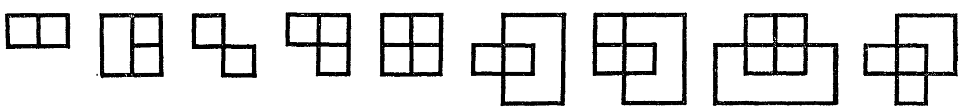


FIG. 4.—Topological patterns for which the monotonicity result is valid and that always lead to exactly one circle.

The prior results of this paper may be easily generalized to other games of a similar nature. For instance, instead of using rectangular arrays of squares, one could use regular arrays of equilateral triangles, regular hexagons, or even cubes. The cells do not even have to have equal numbers of sides. The requirements appear to be simply that there be two types of entities called, say, cells and sides. Then for each cell there are at least two sides associated with it, and for each side there are at most two cells with which it is associated.

Consider, however, structures of the following type, of which the previously cited arrays of squares, triangles or hexagons constitute examples: Let there be given a finite number of points in the plane called dots, and let the sides consist of admissible line intervals between the dots. Let no interval intersect any other interval except at its end points, which are dots. Let all of the finite polygonal components of the plane thereby formed be the cells.

If the sides and dots are connected, then a well-known theorem due to Euler says that (with # meaning "the number of")

$$\# \text{ dots} = \# \text{ sides} - \# \text{ polygonal cells} + 1.$$

Therefore, the total number of moves in the game equals the number of dots plus the number of times that cells are completed as a circle. So if the number of dots is even, the first player should seek an odd number of circles, but if the number of dots is odd, he should seek an even number of circles.

ON COLORING THE $n \times n$ CHESSBOARD

M. R. IYER AND V. V. MENON, Indian Statistical Institute, Calcutta

1. Introduction. We consider the $n \times n$ chessboard; that is, n^2 cells arranged in a square array of n rows and n columns.

DEFINITION 1. *Given a piece (the Queen, say), two cells are called adjacent if the piece can go from one cell to the other in one move.*

Thus one can look upon the chessboard as a graph whose vertices correspond to the cells (on the board) accessible to a given piece, two vertices being joined by an (unoriented) edge if, and only if, the corresponding cells are adjacent.

DEFINITION 2. *For the piece P, the above graph is called the graph of P's move.*

DEFINITION 3. *A coloring of the graph G_p of P's move is an assignment of m colors ($m \geqq 1$) to the vertices of G_p such that two vertices which are joined by an edge do not have the same color. Equivalently, it is the assignment of m colors to the cells of the chessboard such that two adjacent cells do not get the same color.*

For example, if we assign different colors to all the cells, we obtain a coloring.

DEFINITION 4. *Consider the graph G_p of P's move. The chromatic number $k(n; P)$ of G_p is the minimum number of colors required for a coloring of G_p.*

In this paper, we are concerned with the

PROBLEM. Determine the chromatic number of the graph G_p, when P is, respectively, a Knight, a Bishop, a Rook, a Queen, and a King.

The problem is naturally the most interesting when P is the Queen; for this case a partial solution is obtained. The other pieces are included for the sake of completeness.

The coloring of the graph of the Queen for $n=8$ with 9 colors (Fig. 2) was discovered by the first of the authors. She also suggested the Color Scheme 4 with $r=2$ (this is also given in [1]). The generalization of it and the remainder of the paper are due to the second of us.

2. Notations and preliminaries. We name the cells of the chess-board, and hence the corresponding vertices of G_p, as follows: the cell in the ith row and jth column is denoted by the pair (i, j), $1 \leqq i, j \leqq n$.

DEFINITION 5. *Two cells (i, j) and (i', j') are said to be in the same diagonal if $i+j=i'+j'$, or if $i-j=i'-j'$.*

In the following we will refer also to the natural coloring of the chess-board; in this, all the cells (i, j) with $i+j$ an even number are colored white, the remaining being black. Thus the cell (1, 1) is white, and along rows and columns, cells are white and black alternately.

DEFINITION 6. *A clique of the graph G_p is a subset S of vertices such that any two of vertices of S are adjacent to each other (are joined by an edge).*

The following are immediate consequences.

LEMMA 1. *In the graph G_p, consider a clique S. In the coloring of G_p, all vertices of S receive different colors.*

LEMMA 2. *Given a coloring of the graph G_p for the $n \times n$ board ($n \geqq 2$), we obtain a coloring for the $(n-1) \times (n-1)$ board (for the same P) by taking only the $(n-1)^2$ cells in the first $(n-1)$ rows and the first $(n-1)$ columns.*

3. Results. In each of the following cases, we shall determine the chromatic numbers as follows: first, we shall determine a clique with k vertices say, and using the Lemma 1 above, we will see that at least k colors will be needed in any coloring; secondly, we will exhibit a coloring with k colors. These two facts will give k as the chromatic number.

The coloring for the graph G_p and the corresponding coloring for the chessboard, these two are used interchangeably in the following, depending upon convenience. Also, we number the colors 0, 1, 2, 3 $\cdots$.

Naturally, the cases of interest are $n \geqq 2$. For, when $n = 1$, the chromatic number is 1 in all cases.

(1) *The graph of Knight's move.*

This graph consists of n^2 vertices, corresponding to the n^2 cells of the board. The cell (i, j) is adjacent to each one of the set

$$S(i,j) = \left\{ (i',j') \middle| \begin{array}{ll} 1 \leqq i' \leqq n, & 1 \leqq j' \leqq n; \\ \text{and either} & i' = i \pm 1, j' = j \pm 2 \\ \text{or} & i' = i \pm 2, j' = j \pm 1 \end{array} \right\}$$

and to no other cell.

Obviously, if $n = 2$, no two cells are adjacent, and therefore just one color is sufficient to color the whole board.

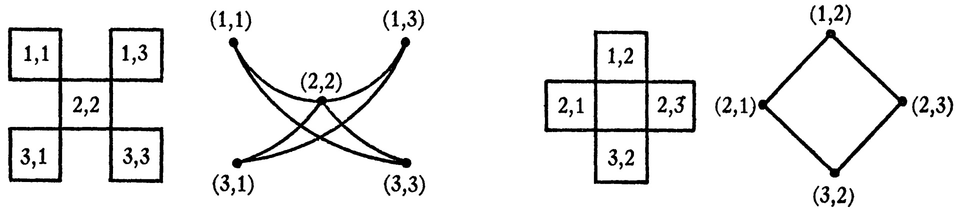
Color Scheme 1. If $n > 2$, at least one pair of adjacent vertices (a clique with two vertices) exists. On the other hand, in the natural coloring, assigning the color $i+j$ (mod 2) to cell (i, j) of the board, two adjacent cells (for the Knight's move) have different colors. Hence the chromatic number is 2.

RESULT 1. $k(n;\text{ knight}) = 1$ if $n = 2$; $= 2$ if $n > 2$.

(2) *The graph of Bishop's move.*

For the White Bishop, the vertices of the graph are the cells (i, j) for which $i+j$ is even, whereas for the Black Bishop, the vertices are the cells (i, j) for which $i+j$ is odd. In either case, two vertices (cells) are adjacent if, and only if, they are on the same diagonal.

The graphs of the bishops are shown in Fig. 1 (a) and (b), for $n = 3$. (The permissible cells on the board are shown on the left with the numbers of the cells, and the graph on the right, in each case.)



(a) White Bishop (b) Black Bishop

FIG. 1

6	2	7	3	1	5	0	4
0	1	5	8	4	7	3	2
5	4	0	1	3	2	6	8
2	7	3	4	6	1	5	0
3	6	2	5	7	0	4	1
7	5	1	0	2	3	8	6
1	0	4	7	8	6	2	3
8	3	6	2	0	4	1	5

FIG. 2

When n is even, the graphs for the two Bishops are the same (isomorphic).

All the cells in a diagonal obviously form a clique. In particular, the cells in the longest diagonal (that containing the largest number of cells) form a clique. Let l be the number of cells in this diagonal. Of course,

$$\begin{aligned} l &= n - 1 \quad \text{for the Black bishop when } n \text{ is odd,} \\ &= n \qquad\quad \text{otherwise.} \end{aligned}$$

We now exhibit a coloring with l colors.

Color Scheme 2. Assign color $i-1$ to cell (i, j), $1 \leqq i \leqq n-1$. The cells in the nth row receive the color 0 if n is odd and the Bishop is Black; in all other cases they receive the color $n-1$. (This scheme colors the board in l colors such that no color repeats along a diagonal.)

Hence l is the chromatic number in this case.

RESULT 2. $k(n;\text{White Bishop}) = n$; $k(n;\text{ Black Bishop}) = n-1$ if n is odd, $=n$ if n is even.

(3) *The graph of Rook's move.*

This consists of n^2 vertices (corresponding to the n^2 cells of the board), two vertices being adjacent if, and only if, the corresponding cells are in the same row or the same column.

The cells in the first row form a clique so that the chromatic number is at least n. On the other hand the following supplies a coloring with n colors.

Color Scheme 3. Assign the color $|i-j|$ to the cell (i, j), $1 \leqq i \leqq j \leqq n$, and the color $j-i+n$ to the cell (i, j), $1 \leqq j < i \leqq n$. Thus, only the cells along a diagonal can have the same color. Therefore, n is the chromatic number for this graph.

RESULT 3. $k(n; \text{Rook}) = n$.

(4) *The graph of Queen's move.*

The vertices of this graph are all the n^2 cells of the board, two vertices being adjacent if, and only if, the corresponding cells are in the same row, same column or the same diagonal.

Clearly the cells in the first row form a clique, so that the chromatic number is at least n.

Color Scheme 4. Assign the colors $0, 1, 2, \cdots, n-1$ to the cells $(1, 1)$, $(1, 2), \cdots, (1, n)$, (along the first row) respectively. Take an integer r, $1 < r < n$. Then the color assigned to the cell $(i+1, j+r \pmod n)$ is the same as that for the cell (i, j), $1 \leqq i < n$, $1 \leqq j \leqq n$.

LEMMA. *The color scheme* 4 *colors the board for the Queen's move if each one of* $r-1$, r, $r+1$ *is prime to* n.

Proof. Take any two cells (i, j) and (i', j'), $i \leqq i'$, having the same color. From the construction it follows that $i < i'$, and $j' = j + (i'-i)r \pmod n$. It is easily verified that $i \neq i'$ (the cells are not in the same row), $j' \neq j$ (the cells are not in the same column), $i+j \neq i'+j'$ and $i-j \neq i'-j'$ (the cells are not in the same diagonal). The last assertions follow because r, $r \pm 1$ are all prime to n.

Hence, if there exists an integer r such that r, $r \pm 1$, are all prime to n, then we can color the board with n colors. Notice, however, that out of the integers r, $r \pm 1$, at least one is divisible by 2 and one by 3. So we can re-state the lemma (with $r = 2$) in the following form. (This result is given in Kraitchik [1] also.)

THEOREM. *If n is not divisible by* 2 *or* 3, *the chromatic number of the graph of Queen's move is n.*

We can now use the Lemma 2 of section 2 for obtaining an upper bound for the chromatic number when n is not of the above form. For example, if n is divisible by 6, then $n+1$ satisfies the conditions for the theorem, so that we can color the $(n+1) \times (n+1)$ board with $n+1$ colors, and hence the $n \times n$ board by $n+1$ colors at most. Now out of the numbers n, $n+1$, $n+2$ and $n+3$, at least one must satisfy the condition of the theorem. Therefore, we obtain

RESULT 4. $k(n; \text{Queen}) = n$ if n is not divisible by 2 or 3; $\leqq n+1$ if $n+1$ is not divisible by 2 or 3; $\leqq n+2$ if n is an odd number divisible by 3; $\leqq n+3$ otherwise.

In all cases, $k(n; \text{Queen}) \geqq n$.

Counterexamples exist to show that this is not the best possible coloring, e.g. $8 \equiv 2 \pmod 3$ requires only 9 colors. See Figure 2. This only gives an upper bound to the number of colors that may be required.

(5) *The graph of King's move.*

This graph consists of all the n^2 cells as vertices; two different cells (i, j) and (i', j') are adjacent if, and only if,

$$|i - i'| = 0 \text{ or } 1, \quad \text{and} \quad |j - j'| = 0 \text{ or } 1.$$

Obviously the four cells (1, 1), (1, 2), (2, 1) and (2, 2) form a clique. Hence $k(n; \text{King}) \geqq 4$.

We can, on the other hand, color the graph with four colors as follows.

Color Scheme 5. Assign color 0 to (1, 1), 1 to (1, 2), 2 to (2, 1) and 3 to (2, 2). Then assign the color of the cell (i, j) to all cells (i, j') with $j \equiv j' \pmod 2$, and to all cells (i', j) with $i \equiv i' \pmod 2$.

RESULT 5. $k(n; \text{King}) = 4$ for $n \geqq 2$.

4. Remarks.

(1) In a graph, an independent set is a subset S of the vertices of the graph, such that no two of the points of S are adjacent. The independence number (see [2]) is defined as the maximum possible number of elements in an independent set, and the corresponding independent set is called a maximal independent set. Scheid [3] has found the independence numbers of the various graphs dealt with in the previous section.

A related problem will be: what is the maximum number of mutually disjoint maximal independent sets in the graph of a piece P?

We notice that in our coloring problem, the vertices which receive a particular color form an independent set. Thus it is easy to give a solution to the above problem (using our color schemes) in the case of Knight, Bishop, Rook and King. For example, one can verify that for a Rook we can have n and no more of mutually disjoint maximal independent sets; our theorem can be restated: "if n is not divisible by 2 or 3, we can find n and no more of mutually disjoint maximal independent sets for the graph of the Queen."

(2) One may verify that k (2; Queen) $=4$, k (3; Queen) $= k$ (4; Queen) $=5$, and k (6; Queen) $=7$.

(3) All the maximal independent sets, in case of Queen, for $n=8$ are well known. If k (8; Queen) $=8$ it will mean that there are 8 mutually disjoint such sets, which we do not expect. The first of us has succeeded in obtaining a coloring with 9 colors. We can conjecture that $k(8; \text{Queen}) = 9$.

(4) It is interesting to note that in the basic construction when n is not a multiple of 2 or 3, the $n \times n$ chess-board is treated as a torus; but this is not the case with the special construction of Figure 2 for $n=8$. In Figure 2 there is a broken diagonal with two 0's and two 2's. It begins with a 0 in column 1 and a 2 in column 2 and continues with 6, 7, 2, 0, 5, 8.

(5) More generally, one can expect that $n \leqq k$ $(n; \text{Queen}) \leqq n+1$ for $n > 3$.

References

1. M. Kraitchik, Mathematical Recreations, Allen and Unwin, London, 1960, pp. 251.
2. O. Ore, Theory of graphs, Amer. Math. Soc. Colloq. Publ., 38 (1962).
3. F. Scheid, Some packing problem, this MONTHLY, 67 (1960) 231–235.

THE WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

J. H. McKAY, Oakland University

The following results of the twenty-sixth William Lowell Putnam Mathematical Competition held on November 20, 1965, have been determined in accordance with the regulations governing the Competition. This competition is supported by the William Lowell Putnam Intercollegiate Memorial Fund left by Mrs. Putnam in memory of her husband and is held under the auspices of the Mathematical Association of America.

The first prize, five hundred dollars, is awarded to the Department of Mathematics of Harvard University, Cambridge, Massachusetts. The members of the team were Daniel Fendel, Roger Howe and Barry Simon; to each of these a prize of fifty dollars is awarded.

The second prize, four hundred dollars, is awarded to the Department of Mathematics of Massachusetts Institute of Technology, Cambridge, Massachusetts. The members of the team were William Ackerman, Michael R. Rolle and Robert Wolf; to each of these a prize of forty dollars is awarded.

The third prize, three hundred dollars, is awarded to the Department of Mathematics of University of Toronto, Toronto, Ontario. The members of the team were Stephen A. Marion, Richard L. Penner and Lon M. Rosen; to each of these a prize of thirty dollars is awarded.

The fourth prize, two hundred dollars, is awarded to the Department of Mathematics of Princeton University, Princeton, New Jersey. The members of the team were James Baker, Daniel Cohen and Robert Weber; to each of these a prize of twenty dollars is awarded.

The fifth prize, one hundred dollars, is awarded to the Department of Mathematics of California Institute of Technology, Pasadena, California. The members of the team were Frederic A. Ferdman, Stacy G. Langton and Vern S. Poythress; to each of these a prize of ten dollars is awarded.

The five persons ranking highest in the examination, named in alphabetical order, are Andreas R. Blass, University of Detroit; Robert Bowen, University of California, Berkeley; Daniel Fendel, Harvard University; Lon M. Rosen, University of Toronto; Barry Simon, Harvard University. To each of these a prize of seventy-five dollars is awarded. The William Lowell Putnam Prize Scholarship at Harvard has been awarded to Mr. Andreas R. Blass, University of Detroit. The value of this scholarship is $2500 plus tuition at Harvard.

The six persons ranking second highest in the examination, named in alphabetical order, are Roger Howe, Harvard University; Neal I. Koblitz, Harvard University; Michael R. Rolle, Massachusetts Institute of Technology; Richard P. Stanley, California Institute of Technology; Charles R. Zarnke, University of Waterloo; and Derek A. Zave, Case Institute of Technology.

The following teams, named in alphabetical order, won honorable mention: Brown University, Providence, Rhode Island, the members of the team being Roger M. Firestone, John R. Hall, Jr.

and Carl B. Pomerance; University of Colorado, Boulder, Colorado, the members of the team being Charles N. Johnson, Theodore S. Martner and Paul R. Sparks; University of Illinois, Urbana, Illinois, the members of the team being Paul H. Cox, Hugh L. Montgomery and Stanislav L. Vrscaj; Michigan State University, East Lansing, Michigan, the members of the team being Allen J. Beadle, William A. Webb and Neil L. White; Swarthmore College, Swarthmore, Pennsylvania, the members of the team being Mayson Lancaster, Robert MacPherson and Stephen B. Maurer.

Honorable mention is given to the following twenty-four individuals, named in alphabetical order: Frank Bernhart, University of Oklahoma; Fred L. Bookstein, University of Michigan; Charles Brenner, University of California, Los Angeles; Marshall Buck, Harvard University; Douglas Burke, Harvard University; Sylvain Cappell, Columbia University; Teddy Chang, Massachusetts Institute of Technology; David Chu, California Institute of Technology; Glenn E. Engebretsen, California Institute of Technology; Irwin Gaines, Harvard University; Kim K. Gibson, California Institute of Technology; David Haynor, Harvard University; B. L. Keyfitz, University of Toronto; Robert MacPherson, Swarthmore College; Sidney Marshall, Dartmouth College; Theodore S. Martner, University of Colorado; Hugh L. Montgomery, University of Illinois; Carl B. Pomerance, Brown University; Vern S. Poythress, California Institute of Technology; Gary Russell, University of Rochester; Neil L. White, Michigan State University; Terry A. Winograd, Colorado College; and Robert Wolf, Massachusetts Institute of Technology.

A total of two thousand sixty-seven contestants from two hundred forty-seven colleges and universities entered the Competition. One thousand five hundred ninety-six contestants from two hundred thirty-five colleges and universities participated in the examination on November 20, 1965.

A listing of the top five hundred contestants may be obtained from the Director. The list, which includes addresses and expected dates of graduation, may be helpful to departments of mathematics in selecting graduate students.

Those participating in the competition were given the following problems to solve:

Part A

A-1. Let ABC be a triangle with angle $A <$angle $C < 90° <$angle B. Consider the bisectors of the external angles at A and B, each measured from the vertex to the opposite side (extended). Suppose both of these line-segments are equal to AB. Compute the angle A.

A-2. Show that, for any positive integer n,

$$\sum_{r=0}^{[(n-1)/2]} \left\{\frac{n-2r}{n}\binom{n}{r}\right\}^2 = \frac{1}{n}\binom{2n-2}{n-1},$$

where $[x]$ means the greatest integer not exceeding x, and $\binom{n}{r}$ is the binomial coefficient "n choose r," with the convention $\binom{n}{0}=1$.

A-3. Show that, for any sequence $a_1, a_2, \cdots$ of real numbers, the two conditions

(A) $$\lim_{n\to\infty} \frac{e^{(ia_1)} + e^{(ia_2)} + \cdots + e^{(ia_n)}}{n} = \alpha$$

and

(B) $$\lim_{n\to\infty} \frac{e^{(ia_1)} + e^{(ia_4)} + \cdots + e^{(ia_{n^2})}}{n^2} = \alpha$$

are equivalent.

A-4. At a party, assume that no boy dances with every girl but each girl dances with at least one boy. Prove that there are two couples gb and $g'b'$ which dance whereas b does not dance with g' nor does g dance with b'.

A-5. In how many ways can the integers from 1 to n be ordered subject to the condition that, except for the first integer on the left, every integer differs by 1 from some integer to the left of it?

A-6. In the plane with orthogonal Cartesian coordinates x and y, prove that the line whose equation is $ux+vy=1$ will be tangent to the curve $x^m+y^m=1$ (where $m>1$) if and only if $u^n+v^n=1$ and $m^{-1}+n^{-1}=1$.

Part B

B-1. Evaluate

$$\lim_{n\to\infty}\int_0^1\int_0^1\cdots\int_0^1 \cos^2\left\{\frac{\pi}{2n}(x_1+x_2+\cdots x_n)\right\}dx_1dx_2\cdots dx_n.$$

B-2. In a round-robin tournament with n players $P_1, P_2, \cdots, P_n$ (where $n>1$), each player plays one game with each of the other players and the rules are such that no ties can occur. Let ω_r and l_r be the number of games won and lost, respectively, by P_r. Show that

$$\sum_{r=1}^{n} w_r^2 = \sum_{r=1}^{n} l_r^2.$$

B-3. Prove that there are exactly three right-angled triangles whose sides are integers while the area is numerically equal to twice the perimeter.

B-4. Consider the function

$$f(x,n)=\frac{\binom{n}{0}+\binom{n}{2}x+\binom{n}{4}x^2+\cdots}{\binom{n}{1}+\binom{n}{3}x+\binom{n}{5}x^2+\cdots},$$

where n is a positive integer. Express $f(x, n+1)$ rationally in terms of $f(x, n)$ and x. Hence, or otherwise, evaluate $\lim_{n\to\infty} f(x, n)$ for suitable fixed values of x. (The symbols $\binom{n}{r}$ represent the binomial coefficients.)

B-5. Consider collections of unordered pairs of V different objects $a, b, c, \cdots, k$. Three pairs such as bc, ca, ab are said to form a triangle. Prove that, if $4E\leqq V^2$, it is possible to choose E pairs so that no triangle is formed.

B-6. If A, B, C, D are four distinct points such that every circle through A and B intersects (or coincides with) every circle through C and D, prove that the four points are either collinear (all of one line) or concyclic (all on one circle).

Solutions. Part A

A-1. Suppose the bisector of the exterior angle at A intersects line BC at X and the bisector of the exterior angle at B meets the line AC at Y. The assumption that C is between B and X contradicts the fact that $\angle B>\angle C$ so we may assume that B is between X and C. Similarly, we conclude that C is between A and Y because $\angle A<\angle C$.

If Z is a point on line AB with B between A and Z, we have from triangle ABY that $\angle ZBY=2A$. Hence, $\angle BXA=\angle ABX=\angle ZBC=2\angle ZBY=4A$, and the angle sum of triangle ABX is $90°-\frac{1}{2}A+8A$. Thus, $A=12°$.

Comment: The only thing that takes this problem out of the category of a simple high school exercise is the "justification of the figure" or more precisely the consideration of all cases. Most contestants who essentially solved the problem were very lax in explicitly using the hypotheses to justify the betweenness relations which they implicitly assumed.

A-2. Substituting $s=n-r$ in the given summation reveals that twice this sum is equal to:

$$\sum_{r=0}^{n}\left\{\frac{n-2r}{n}\binom{n}{r}\right\}^2$$

$$=\sum\left(1-2\frac{r}{n}\right)^2\binom{n}{r}^2=\sum\binom{n}{r}^2-4\sum\frac{r}{n}\binom{n}{r}\binom{n}{r}+4\sum\left(\frac{r}{n}\right)^2\binom{n}{r}^2$$

$$=\binom{2n}{n}-4\sum\binom{n-1}{r-1}\binom{n}{r}+4\sum\binom{n-1}{r-1}^2$$

$$=\binom{2n}{n}-4\binom{2n-1}{n-1}+4\binom{2n-2}{n-1}=\binom{2n}{n}-4\binom{2n-2}{n-2}$$

$$=\left\{\frac{2n(2n-1)}{n^2}-4\frac{n-1}{n}\right\}\binom{2n-2}{n-1}=\frac{2}{n}\binom{2n-2}{n-1}.$$

Comment: This solution assumes the well-known identities

$$\sum\binom{n}{r}^2=\binom{2n}{n}\quad\text{and}\quad\sum_{r=0}^{k}\binom{m}{k-r}\binom{n}{r}=\binom{m+n}{k}$$

which may be proved by comparing coefficients in the expansion of

$$(1+x)^m\cdot(1+x)^n=(1+x)^{m+n}.$$

A-3. That (A) implies (B) follows from the fact that subsequences of a convergent sequence converge to the limit of the sequence. We simplify the notation by setting $c_r=\exp ia_r$ and $S(t)=c_1+c_2+\cdots+c_t$. Note that $|c_r|=1$ and $|S(t+k)-S(t)|\leqq k$. Suppose now that (B) holds and write $m=n^2+k$, where $0\leqq k\leqq 2n$.

$$\left|\frac{S(m)}{m}-\frac{S(n^2)}{n^2}\right|\leqq\left|\frac{S(m)}{m}-\frac{S(n^2)}{m}\right|+\left|\frac{S(n^2)}{n^2}-\frac{S(n^2)}{m}\right|$$

$$\leqq\frac{k}{m}+n^2\left(\frac{1}{n^2}-\frac{1}{m}\right)=\frac{k+m-n^2}{m}=\frac{2k}{m}\leqq\frac{4n}{n^2}.$$

We conclude that $\lim_{m\to\infty}(S(m)/m-S(n^2)/n^2)=0$ or that $S(m)/m$ converges to α.

A-4. Let b be a boy who dances with a maximal number of girls (i.e., there may be another boy who dances with the same number of girls, but none dances

with a greater number). Let g' be a girl with whom b does not dance, and b' a boy with whom g' dances. Among the partners of b, there must be at least one girl g who does not dance with b' (for otherwise b' would have more partners than b). The couples gb and $g'b'$ solve the problem.

Comment: The contestants submitted at least four solutions which were more or less distinct. Some were in graph theoretic terms and some in matrix form.

A-5. On the basis of the first few cases we conjecture the answer is 2^{n-1} and proceed by induction.

We first show (also by induction) that an n-arrangement ends in 1 or n. Note that when n is deleted from an n-arrangement, the result is an $(n-1)$-arrangement. If an n-arrangement does not end in 1 or n, deletion of n produces an $(n-1)$-arrangement ending (by induction) in $(n-1)$. This implies the n-arrangement ended in n because n cannot precede $(n-1)$.

For any n-arrangement $(a_1, a_2, \cdots, a_n)$ there is another n-arrangement $(a_1', a_2', \cdots, a_n')$, where $a_i' = n+1-a_i$. If one of these ends in n, the other ends in 1 and consequently exactly half of the n-arrangements end in n.

All of the n-arrangements which end in n can be obtained by adjoining an n to the end of all $(n-1)$-arrangements, and by induction there are 2^{n-2} of these. Hence, there are 2^{n-1} n-arrangements.

Comment: Several counting techniques were used by the contestants and many were quite ingenious.

A-6. The problem is not well set being true only under rather heavy restrictions on the x, y, u and v. For example, all is in order if they are nonnegative. However, if m is rational with odd numerator and odd denominator there are tangent lines to the curve for which $u^n+v^n>1$, while if m is rational with odd numerator and even denominator then n is rational with odd numerator and odd denominator and there are solutions (u, v) of $u^n+v^n=1$ such that the line $ux+vy=1$ is not tangent to the curve $x^m+y^m=1$.

No contestant, however, noted any of these facts explicitly and major credit was given for a solution covering the case where x, y, u, v are positive.

Let (x_0, y_0) be a point on the curve $x^m+y^m=1$. The tangent to this curve is $x_0^{m-1}x+y_0^{m-1}y=1$. If this line is $ux+vy=1$, then $u=x_0^{m-1}$ and $v=y_0^{m-1}$ with both u and v nonnegative. The relation $1/m+1/n=1$ gives $m/(m-1)=n$ and we obtain $u^n+v^n=x_0^m+y_0^m=1$.

Conversely, let $m^{-1}+n^{-1}=1$ and let u and v be nonnegative and such that $u^n+v^n=1$. Define x_0 and y_0 by the equations $x_0=u^{n/m}$ and $y_0=v^{n/m}$. Then x_0 and y_0 are nonnegative and $x_0^m+y_0^m=u^n+v^n=1$. Thus, (x_0, y_0) is on the curve $x^m+y^m=1$ and the line $ux+vy=1$ is the tangent to the curve by the calculation above.

In this solution we use the fact that for nonnegative a and positive r and s $(a^r)^s=a^{rs}$.

Solutions. Part B

B-1. The change of variables $x_k \to 1 - x_k$ yields

$$\int_0^1\int_0^1 \cdots \int_0^1 \cos^2\left\{\frac{\pi}{2n}(x_1 + x_2 + \cdots + x_n)\right\} dx_1 dx_2 \cdots dx_n$$
$$= \int_0^1\int_0^1 \cdots \int_0^1 \sin^2\left\{\frac{\pi}{2n}(x_1 + x_2 + \cdots + x_n)\right\} dx_1 dx_2 \cdots dx_n.$$

Each of these expressions, being equal to half their sum, must equal $\frac{1}{2}$. The limit is also $\frac{1}{2}$.

B-2. Clearly $\omega_r + l_r = n - 1$ for $r = 1, 2, \cdots, n$ and $\sum_1^n \omega_r = \sum_1^n l_r$. Hence,

$$\sum_1^n \omega_r^2 - \sum_1^n l_r^2 = \sum_1^n (\omega_r - l_r)(\omega_r + l_r) = (n-1)\sum_1^n (\omega_r - l_r) = (n-1)\cdot 0 = 0.$$

B-3. All Pythagorean triples can be obtained from $x = \lambda(p^2 - q^2)$, $y = 2\lambda pq$, $z = \lambda(p^2 + q^2)$ where $0 < q < p$, $(p, q) = 1$ and $p \not\equiv q \bmod 2$, λ being any natural number.

The problem requires that $\frac{1}{2}xy = 2(x + y + z)$. This condition can be written $\lambda^2(p^2 - q^2)(pq) = 2\lambda(p^2 - q^2 + 2pq + p^2 + q^2)$ or simply $\lambda(p - q)q = 4$. Since $p - q$ is odd it follows that $p - q = 1$ and the only possibilities for q are 1, 2, 4.

If $q = 1$, $p = 2$, $\lambda = 4$, $x = 12$, $y = 16$, $z = 20$.
If $q = 2$, $p = 3$, $\lambda = 2$, $x = 10$, $y = 24$, $z = 26$.
If $q = 4$, $p = 5$, $\lambda = 1$, $x = 9$, $y = 40$, $z = 41$.

Comment: Many contestants recalled the form of the primitive Pythagorean triples, which omits the λ in the above expressions for x, y, and z. For this particular problem, they were led to the correct triangles but the reasoning is clearly incomplete. Somewhat more than half credit was allowed for this effort.

The most satisfying solutions were those which proceeded from first principles foregoing any reliance on the known formula for Pythagorean triples.

B-4. Since

$$\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}, \qquad f(x, n+1) = \frac{f(x, n) + x}{f(x, n) + 1}.$$

If x is such that $f(x, n)$ converges when n tends to infinity, the limit $F(x)$ must satisfy $F(x) = (F(x) + x)/(F(x) + 1)$, $F^2(x) = x$. The convergence to $\sqrt{x}$ is obvious when $x = 0$ or 1. To show this convergence for any positive x we first note that

$$f(x, n) = \sqrt{x}\,\frac{(1 + \sqrt{x})^n + (1 - \sqrt{x})^n}{(1 + \sqrt{x})^n - (1 - \sqrt{x})^n}.$$

When $0<x<1$, write $a=(1-\sqrt{x})/(1+\sqrt{x})$; then $0<a<1$ and

$$f(x, n) = \sqrt{x}\frac{1+a^n}{1-a^n} \to \sqrt{x}.$$

When $x>1$, write $b=(\sqrt{x}-1)/(\sqrt{x}+1)$; then $0<b<1$ and

$$f(x, n) = \sqrt{x}\frac{1+(-b)^n}{1-(-b)^n} \to \sqrt{x}.$$

The limit fails to exist for negative values of x; but for all other complex numbers the limit exists and is that square root of x which lies in the right half plane.

Comment: Full credit was given for a complete analysis of the real case.

B-5. Divide the objects into two subsets $\{a_1, a_2, \cdots, a_m\}$ and $\{b_1, b_2, \cdots, b_n\}$, where $m+n=V$. Then the mn pairs (a_j, b_k), where $j=1, 2, \cdots, m$ and $k=1, 2, \cdots, n$, obviously contain no triangles. If V is even, take $m=n=V/2$, and if V is odd, take $m=(V+1)/2$, $n=(V-1)/2$. Then $mn \geqq V^2/4 \geqq E$.

B-6. Suppose A, B, C, D are neither concyclic nor collinear. Then p, the perpendicular bisector of segment AB, cannot coincide with q, the perpendicular bisector of segment CD. If the lines p and q intersect, their common point is the center of two *concentric* circles, one through A and B, the other through C and D. If instead p and q are parallel, so also are the lines AB and CD. Consider points P and Q, on p and q respectively, midway between the parallel lines AB and CD. Clearly, the circles ABP and CDQ have no common point.

Comment: This was a difficult problem to grade. There were many awkward but essentially valid solutions presented and many more that, despite much circumlocution, failed to be convincing.

Acknowledgement

The Director would like to acknowledge the assistance of H. S. M. Coxeter and L. M. Kelly in drafting the above solutions and acknowledge the service of the following persons, who were graders for the competition: H. R. Atkinson, J. M. Calloway, D. W. Crowe, J. W. Dettman, J. A. Eagon, R. A. Gambill, S. I. Goldberg, N. D. Kazarinoff, L. M. Kelly, M. S. Klamkin, E. B. Leach, W. S. Loud, E. A. Nordhaus, E. T. Wong.

ME

I am not, as a general rule, averse to Work,
but I do, I fear, have Moments of Inertia.
I go around in circles with a Torque,
and in beyond my depth with Liquid Pressure.
While the bulk o' me is levity,
I have a Center of Gravity.

KATHARINE O'BRIEN

EDITED BY J. H. CURTISS, University of Miami

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457

ON THE UNIFORM LIMIT OF MULTIPLE BALAYAGE OF VECTOR INTEGRALS

HUBERT HALKIN, University of California at San Diego

Introduction. Let f be an integrable function from $[0, 1]$ into an n-dimensional Euclidean space. The vector $I=\int_0^1 f d\mu$ represents the average of the function f on the interval $[0, 1]$. The vector I is also the value at $t=1$ of the function $g(t)=\int_0^t f d\mu$. If we consider the estimation of the average I as a continuous process then the vector $g(t)$ may be regarded as a certain approximation of the fraction tI of the average vector I. This continuous estimation process is not very accurate if the function f fluctuates greatly. Instead of basing our estimation on a single balayage of the interval $[0, 1]$ we could consider a simultaneous balayage of each of the intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ and introduce a function $g_1(t)$ defined over $[0, 1]$ by the relation:

$$g_1(t) = \int_0^{t/2} f d\mu + \int_{1/2}^{1/2+t/2} f d\mu$$

as an approximation of the fraction tI of the average vector I. Intuitively we could hope that the fluctuations of the function f on the intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ will compensate each other and that $g_1(t)$ will be a better approximation of tI than $g(t)$. This process may be refined further: for each integer k we partition the interval $[0, 1]$ into 2^k consecutive intervals of equal length $1/2^k$ and we consider the function $g_k(t)$ defined over $[0, 1]$ by the relation:

$$g_k(t) = \sum_{i=1}^{2^k} \int_{(i-1)/2^k}^{(i-1+t)/2^k} f d\mu$$

as the kth order approximation of the fraction tI of the average value I. In this paper we prove that these estimations are becoming more and more accurate as k increases. More precisely we shall prove the following result:

THEOREM I. *The function $g_k(t)$ converges uniformly to the function tI as k tends to infinity.*

Theorem I is closely related to Lyapounov's Theorem on the range of a vector integral [**1**, **2**, **5**, **6**, **7**] and has some useful applications in the theory of optimal control [**3**, **4**].

Proof of Theorem I. We shall first assume that the integrable function f is real valued and bounded. We have then $f\in\mathcal{L}^2([0, 1])$ and there exists a sequence of real numbers $a_0, a_1, a_{-1}, a_2, a_{-2}, \cdots$ such that

$$f(t) \approx a_0 + \sum_{i=1}^{\infty} (a_i \cos 2i\pi t + a_{-i} \sin 2i\pi t),$$

where "$\approx$" stands for $\mathcal{L}^2$ convergence, and

$$a_0 = \int_0^1 f(t)dt.$$

Let T be a linear operator from $\mathcal{L}^2([0, 1])$ into $\mathcal{L}^2([0, 1])$ defined by

$$(Th)(t) = \frac{1}{2}\left(h\left(\frac{t}{2}\right) + h\left(\frac{1}{2} + \frac{t}{2}\right)\right).$$

We have then

$$(T^k f)(t) \approx a_0 + \sum_{n=1}^{\infty} (a_{n2^k} \cos 2n\pi t + a_{-n2^k} \sin 2n\pi t),$$

which implies that

$$\lim_{k\to\infty}\left(\int_0^1 |(T^k f)(t) - a_0|^2 dt\right)^{1/2} = 0.$$

It is a trivial matter to verify that $g_k(t) = \int_0^t (T^k f)(\tau)d\tau$. We have then

$$|g_k(t) - a_0 t| \leqq \int_0^1 |(T^k f)(\tau) - a_0| d\tau.$$

By Hölder's inequality we obtain

$$\int_0^1 |(T^k f)(\tau) - a_0| d\tau \leqq \left(\int_0^1 |(T^k f)(\tau) - a_0|^2 d\tau\right)^{1/2}.$$

We have already proved that the right side of the previous inequality tends to 0 as k tends to infinity. This concludes the proof of Theorem I in the case of a real valued bounded integrable function f.

We shall now prove Theorem I in the case of a real valued integrable function f which is not necessarily bounded.

For every positive integer n let

$$S(n) = \{t: t \in [0, 1], |f(t)| > n\},$$

and let f_n be defined over $[0, 1]$ by the relations

$$\begin{aligned} f_n(t) &= 0 \qquad \text{if } |f(t)| > n \text{ and } \mu(S(n)) = 0 \\ &= \frac{\int_{S(n)} f(\tau)d\tau}{\mu(S(n))} \qquad \text{if } \mu(S(n)) \neq 0 \text{ and } |f(t)| > n \\ &= f(t) \qquad \text{if } |f(t)| \leqq n. \end{aligned}$$

We have then $I = \int_0^1 f(t)dt = \int_0^1 f_n(t)dt$ for $n = 1, 2, 3, \cdots$, and

$$\left| g_k(t) - It \right| \leqq \int_{S(n)} |f(\tau)| \, d\tau + \left| \sum_{i=1}^{2k} \int_{(i-1)/2^k}^{(i-1+t)/2^k} f_n(\tau) d\tau - It \right|.$$

The first term on the right side of the previous inequality can be made arbitrarily small by taking n large enough. For a fixed n the second term tends uniformly to zero as k tends to infinity since we have already proved Theorem I for a bounded real valued integrable function. It follows then that $g_k(t)$ tends uniformly to It as k tends to infinity. This concludes the proof of Theorem I in the case of a real valued integrable function f.

In the case of an n-dimensional vector valued integrable function f the previous result can be applied to each component. Hence each component of the sequence of vector valued functions $g_1(t)$, $g_2(t)$, $\cdots$ converges uniformly to the corresponding component of the vector valued function tI. Since the number of components is finite it follows that the sequence of vector valued functions $g_1(t)$, $g_2(t)$, $\cdots$ converges uniformly to the vector valued function tI. This concludes the Proof of Theorem I.

Final remarks. It is also possible to prove Theorem I with the help of the Martingale Theorem, a result in the theory of stochastic processes. The proof given here, however, is based upon mathematical tools, namely elementary results in Fourier Analysis, which are simpler and more widely known than Martingale theory. I am very grateful to Professor D. L. Burkholder for having pointed out the relation between Martingale theory and the present paper, and to Dr. W. L. Roach, Jr., for his valuable help in analysing this relation.

Research sponsored by the AFO of Scientific Research Office of Aerospace Research, US Air Force under AFOSR Grant No. 1039-66.

References

1. D. Blackwell, The range of a vector measure, Proc. Amer. Math. Soc., 2 (1951) 390–395.
2. A. Dvoretzky, A. Wald, and J. Wolfowitz, Relations among certain ranges of vector measures, Pacific J. Math., 1 (1951) 59–74.
3. H. Halkin, Lyapounov's theorem on the range of a vector measure and Pontryagin's maximum principle, Archive for Rational Mechanics and Analysis, 10 (1962) 296–304.
4. ———, On the necessary condition for optimal control of nonlinear systems, J. Analyse Math., 12 (1964) 1–82.
5. P. R. Halmos, The range of a vector measure, Bull. Amer. Math. Soc., 54 (1948) 416–421.
6. S. Karlin, Extreme points of vector functions, Proc. Amer. Math. Soc., 4 (1953) 603–610.
7. A. Lyapounov, Sur les fonctions vecteurs complètement additives, Bull. Acad. Sci. USSR, 4 (1940) 465–478. (In Russian with a French résumé.)

COMMUTING MAPPINGS AND COMMON FIXED POINTS

GERALD JUNGCK, Bradley University

As stated in [1], the question as to whether or not two continuous commuting mappings of the unit interval I into itself have a common fixed point remains unanswered except in certain special cases. Our intent is to offer some

new special cases which may be of interest to those attacking the problem.

To avoid repetition, we assume throughout this note that f and g are continuous mappings of I into itself such that $f(g(x))=g(f(x))$ for $x\in I$. Since our approach focuses attention on the points of intersection of f and g, the following lemma is necessary to guarantee that such points do exist.

LEMMA 1. *The set* $A=\{x\in I: f(x)=g(x)\}$ *is not empty.*

Proof. If A is empty, the continuity of f and g permits us to assume without loss of generality that (1) $f(x)>g(x)$ for $x\in I$. Since $g(0)\geqq 0$, the set $S=\{x\in I: g(x)\geqq x\}$ is not empty. Thus, since S is closed, S has a maximum element c. Clearly, $c=g(c)$. Hence $f(c)=f(g(c))=g(f(c))$, so that $f(c)\in S$. Consequently $f(c)\leqq c=g(c)$. Since $f(c)\leqq g(c)$ contradicts (1), the assumption that A is empty is false, proving the lemma.

Now suppose that $x\in A$. Then $f(x)=g(x)$ and, hence, $f(f(x))=f(g(x))=g(f(x))$ and $g(f(x))=f(g(x))=g(g(x))$. Thus if $x\in A$, $f(x)$, $g(x)\in A$. The proof of the following lemma appeals to this fact.

LEMMA 2. *If f and g have no common fixed point, there exist a, $b\in I$ such that*
(a) $f(a)=g(a)\geqq b>a\geqq f(b)=g(b)$,
(b) $g(x)\neq f(x)$ *for* $x\in(a, b)$.

Proof. The set A of Lemma 1 is closed and hence has a minimum element and a maximum element which we denote by c and d respectively. Since $c\in A$, $f(c)\in A$ so that $f(c)\geqq c$; i.e., $g(c)=f(c)\geqq c$. But c is not a common fixed point by hypothesis; therefore

$$f(c)=g(c)>c. \tag{1}$$

Like reasoning yields

$$f(d)=g(d)<d. \tag{2}$$

Now (1) asserts that the set $S=\{x\in I: f(x)=g(x)\geqq x\}$ is not empty. Moreover, the continuity of f and g assures us that S is closed. Consequently S has a maximum element a ($<d$), where $f(a)=g(a)\geqq a$. As above we conclude that

$$f(a)=g(a)>a. \tag{3}$$

In view of (2), similar reasoning yields a minimum point $b\in[a, d]$ such that

$$f(b)=g(b)<b. \tag{4}$$

Clearly $a<b$.

The designations of a and b force the conclusion that $A\cap(a, b)=\varnothing$, and (b) of the lemma is proved. Since (4) asserts that $f(b)$ is an element of A less than b, $A\cap(a, b)=\varnothing$ implies that $f(b)\leqq a$. Similarly, (3) implies that $f(a)\geqq b$. Thus $f(a)\geqq b>a\geqq f(b)$ so that (3) and (4) imply the conclusion (a) and the lemma is proved.

One result proved by DeMarr in [1] is that f and g have a common fixed

point provided $|f(x)-f(y)| \leqq |x-y|$ for $x, y \in I$. The following theorem yields a less stringent restriction of similar form. We should note that in the proof of this theorem we use the fact that $f(a)$ is a fixed point of g whenever a is a fixed point of g. (This fact was actually verified in the proof of Lemma 1.)

THEOREM 1. *f and g have a common fixed point provided there exists a real number $\alpha > 0$ such that*

$$|f(x) - f(y)| \leqq \alpha |gf(x) - gf(y)| + |x - y|, \qquad \textit{for } x, y \in I.$$

Proof. Suppose that f and g have no common fixed point. Lemma 2 then yields points $a, b \in I$ such that:

(a) $f(a) = g(a) \geqq b > a \geqq f(b) = g(b)$, and

(b) $f(x) \neq g(x)$ for $x \in (a, b)$.

As a consequence of continuity, (b), and the symmetry in (a), we can assume without loss of generality that $g(x) > f(x)$ for $x \in (a, b)$. Since $g(a) \geqq b > a \geqq g(b)$, g has a fixed point and hence a minimum fixed point $d \in (a, b)$. But $g(d) = d$ implies that $f(d)$ is a fixed point of g. Since $f(d) < g(d) = d$, the minimality of d and property (a) demand that $f(d) < a < f(a)$. Thus $\exists x \in (a, d)$ such that $f(x) = a$, and hence a maximum element $c \in (a, d)$ such that $f(c) = a$. It follows that (1) $f(x) < a$ for $x \in (c, d]$.

Now since $f(c) = a$, $g(f(c)) = g(a) = f(a)$; and since

$$(*) \qquad f(a) \geqq b > d$$

we conclude that $gf(c) > d$. On the other hand, $g(d) = d$ implies that $g(f(d)) = f(d)$; hence $gf(d) < d$ since $f(d) < a < d$ (see (1)). We thus have $gf(c) > d > gf(d)$, and we can pick a point $y \in (c, d)$ such that (2) $gf(y) = d$. Moreover, the inequality $f(c) = a < d < f(a)$, see (*), yields a point $x \in (a, c)$ such that $f(x) = d$. Hence (3) $gf(x) = g(d) = d$.

We now complete the proof. First note that since $y \in (c, d)$, (1) demands that $f(y) < a$, and since $f(x) = d$ we can write $f(x) - f(y) > d - a$. But $a < x < y < d$ implies that $(d-a) > (y-x) > 0$, so that $|f(x) - f(y)| > |x - y|$. Since $gf(x) = gf(y)$ by (2) and (3),

$$|f(x) - f(y)| > \alpha |gf(x) - gf(y)| + |x - y|$$

for any positive real number α. We thus have the anticipated contradiction.

We conclude with two results by-passed in the process of proving Theorem 1. First consider the points a and d. We have:

$$f(a) = g(a) > g(d) = d > a > f(d),$$

so that $|f(a) - f(d)| > |g(a) - g(d)| + |a - d|$. Since this inequality is the consequence of the assumption that f and g have no common fixed point, we have proved:

THEOREM 2. *If $|f(x) - f(y)| \leqq |g(x) - g(y)| + |x - y|$ for $x, y \in I$ such that $(f(x) - g(x))(f(y) - g(y)) = 0$, then f and g have a common fixed point.*

Secondly, consider (3) in the proof of Theorem 1. We have $gf(x)=d$, where $f(x)=d$ and $x\in(a,c)$. Thus $f(x)=gf(x)>x$, since $x<c<d$. But $x\in(a,c)\subset(a,b)$ implies that $g(x)>f(x)$, which yields

$$g(x) > f(x) = gf(x) > x.$$

This inequality permits our final assertion:

THEOREM 3. *If there exists a positive real number α such that*

$$\left| x - g(x) \right| \leqq \alpha \left| gf(x) - f(x) \right| + \left| x - f(x) \right|$$

for $x\in I$, f and g have a common fixed point.

Reference

1. Ralph DeMarr, A common fixed point theorem for commuting mappings, this MONTHLY, 70 (1963) 535–537.

ON A CONJECTURE BY THORP

C. W. BURRILL, New York University

1. Introduction. In [1] Thorp obtained two expressions for the probability of a certain event. With a change of notation these were

$$(1)\quad g_m(x_1, \cdots, x_m) = x_0\left[\frac{1}{x_0} - \sum_{k_1\leqq m}\frac{1}{x_0+x_{k_1}} + \sum_{k_1<k_2\leqq m}\frac{1}{x_0+x_{k_1}+x_{k_2}} - \cdots\right]$$

and

$$(2)\quad \begin{aligned} &f_m(x_1, \cdots, x_m) \\ &\quad = x_1\cdots x_m\sum_s\frac{1}{(x_0+x_{s(1)})(x_0+x_{s(1)}+x_{s(2)})\cdots(x_0+\cdots+x_{s(m)})}, \end{aligned}$$

where the sum is over all permutations of the integers 1 to m and where $0\leqq x_k$ $(k=0,\cdots,m)$ and $x_0+\cdots+x_m\leqq 1$. Thorp proved the equality of (1) and (2) in case $x_1=\cdots=x_m$ and conjectured that

$$(3)\quad \frac{m!x_1\cdots x_m}{(x_0+\mu)\cdots(x_0+m\mu)} \leqq f_m(x_1,\cdots,x_m) \leqq \frac{m!\mu^m}{(x_0+\mu)\cdots(x_0+m\mu)},$$

where $m\mu=x_1+\cdots+x_m$. It is the purpose of this note to prove that (1) and (2) are equal in the general case and to prove (3).

2. Equality of f_m and g_m. The basic tool we shall use is the fact that

$$(4)\quad f_m(x_1,\cdots,x_m) = \frac{1}{x_0+\cdots+x_m}\sum_{k=1}^m x_k f_{m-1}(x_1,\cdots,x_{k-1},x_{k+1},\cdots,x_m).$$

To verify this, observe that the right member equals

$$(5)\qquad \frac{x_1 \cdots x_m}{x_0 + \cdots + x_m} \sum_{k=1}^{m} \sum_{\sigma} \frac{1}{(x_0 + x_{\sigma(1)}) \cdots (x_0 + x_{\sigma(1)} + \cdots + x_{\sigma(m-1)})},$$

where the second sum is over all permutations of the integers $1, \cdots, k-1, k+1, \cdots, m$. This may be written

$$(6)\qquad x_1 \cdots x_m \sum_{k=1}^{m} \sum_{\sigma} \frac{1}{(x_0 + x_{\sigma(1)}) \cdots (x_0 + \cdots + x_{\sigma(m)})},$$

where $\sigma(m) = k$. The validity of (4) follows from the observation that each term of (2) appears once and only once as a term of (6).

Next we prove that g_m satisfies a similar equation. To this end, consider

$$\frac{1}{x_0 + \cdots + x_m} \sum_{k=1}^{m} x_k g_{m-1}(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_m)$$

$$(7)\qquad = \frac{x_0}{x_0 + \cdots + x_m} \sum_{k=1}^{m} x_k \sum_{j=0}^{m-1} (-1)^j \sum_{k_1 < \cdots < k_j \leq m, k_i \neq k} \frac{1}{x_0 + x_{k_1} + \cdots + x_{k_j}}$$

$$= \frac{x_0}{x_0 + \cdots + x_m} \sum_{j=0}^{m-1} (-1)^j \sum_{k_1 < \cdots < k_j \leq m} \frac{x_{k_{j+1}} + \cdots + x_{k_m}}{x_0 + x_{k_1} + \cdots + x_{k_j}},$$

where $k_1, \cdots, k_m$ are all of the integers $1, \cdots, m$. Thus the expression of (7) is equal to

$$x_0 \sum_{j=0}^{m-1} (-1)^j \sum_{k_1 < \cdots < k_j \leq m} \left[\frac{1}{x_0 + x_{k_1} + \cdots + x_{k_j}} - \frac{1}{x_0 + \cdots + x_m} \right]$$

$$(8)\qquad = x_0 \sum_{j=0}^{m-1} (-1)^j \sum_{k_1 < \cdots < k_j \leq m} \frac{1}{x_0 + x_{k_1} + \cdots + x_{k_j}} - \frac{x_0}{x_0 + \cdots + x_m} \sum_{j=0}^{m-1} (-1)^j \binom{m}{j}.$$

Since

$$\sum_{j=0}^{m-1} (-1)^j \binom{m}{j} = (-1)^{m+1},$$

it follows that

$$(9)\qquad g_m(x_1, \cdots, x_m) = \frac{1}{x_0 + \cdots + x_m} \sum_{k=1}^{m} x_k g_{m-1}(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_m).$$

Since $f_1 = g_1$, it follows by induction from (4) and (9) that $f_m = g_m$ for all m.

3. Proof of (3). First let us prove that

$$(10)\qquad \frac{m!}{(x_0 + \mu) \cdots (x_0 + m\mu)} \leq \frac{1}{x_1 \cdots x_m} f_m(x_1, \cdots, x_m) = h_m(x_1, \cdots, x_m),$$

where $m\mu = x_1 + \cdots + x_m$, which will establish the lower bound of (3). This holds trivially for $m=1$; assume it to be true for $m-1$.

From (4) it follows that

$$h_m(x_1, \cdots, x_m) = \frac{1}{x_0 + m\mu} \sum_{k=1}^{m} h_{m-1}(x_1, \cdots, x_{k-1}, x_{k+1}, \cdots, x_m). \tag{11}$$

Hence

$$h_m(x_1, \cdots, x_m) \geqq \frac{1}{x_0 + m\mu} \sum_{k=1}^{m} \phi(x_k), \tag{12}$$

where

$$\phi(x_k) = \frac{(m-1)!}{[x_0 + \mu_k] \cdots [x_0 + (m-1)\mu_k]} \tag{13}$$

and $(m-1)\mu_k = m\mu - x_k$. It is easy to verify that $\phi''(x_k) > 0$, hence ϕ is a strictly convex function in the interval $[0, m\mu]$.

Now suppose that the minimum of $\sum \phi(x_k)$ occurs at $\bar{x}_1 \leqq \cdots \leqq \bar{x}_m$. (Clearly the minimum exists since the function is continuous on a compact set.) If $\bar{x}_1 < \bar{x}_m$, then from the strict convexity of ϕ, $\phi(\bar{x}_1) + \phi(\bar{x}_m) > \phi(x_1^*) + \phi(x_m^*)$, where $x_1^* = x_m^* = \frac{1}{2}(\bar{x}_1 + \bar{x}_m)$. Thus we could obtain a smaller sum by replacing $\bar{x}_1$ and $\bar{x}_m$ with x_1^* and x_m^*. We conclude that at the minimum $\bar{x}_1 = \cdots = \bar{x}_m = \mu$. Hence

$$h_m(x_1, \cdots, x_m) \geqq \frac{m\phi(\mu)}{x_0 + m\mu} \tag{14}$$

which yields (10).

To prove that

$$f_m(x_1, \cdots, x_m) \leqq \frac{m!\mu^m}{(x_0 + \mu) \cdots (x_0 + m\mu)} \tag{15}$$

seems to require more care. Clearly (15) holds for $m=1$; assume it to be true for $m-1$. Then from (4)

$$f_m(x_1, \cdots, x_m) \leqq \frac{1}{x_0 + m\mu} \sum_{k=1}^{m} \psi(x_k), \tag{16}$$

where

$$\psi(x_k) = \frac{(m-1)!x_k\mu_k^{m-1}}{[x_0 + \mu_k] \cdots [x_0 + (m-1)\mu_k]} \tag{17}$$

and $(m-1)\mu_k = m\mu - x_k$. If $x_0 = 0$, $\psi(x_k) = (m-1)!x_k$ which leads immediately to (15). Hereafter we suppose $x_0 > 0$.

Employing logarithmic differentiation we find that $\psi'(x_k)=\psi(x_k)\alpha(x_k)$, where

$$(18)\qquad \alpha(x_k)=\frac{1}{x_k}-\frac{1}{\mu_k}+\frac{1}{m-1}\left[\frac{1}{x_0+\mu_k}+\cdots+\frac{m-1}{x_0+(m-1)\mu_k}\right].$$

Examination of $\alpha'(x_k)$ reveals that it attains its largest values at $x_0=0$, and from this it follows that $\alpha'(x_k)\leqq -1/x_k^2$; hence $\alpha(x_k)$ is strictly decreasing. For small values of x_k, $\alpha(x_k)$ is positive, but it must assume some negative values since $\psi'=\psi\alpha$ and $\psi(0)=\psi(m\mu)=0$. We conclude that α has exactly one zero in $[0, m\mu]$; call this point p. Since $\alpha(\mu)>0$, we observe that $p>\mu$. Note, too, that ψ is increasing in $[0, p]$ and decreasing in $[p, m\mu]$.

Next, consider $\psi''(x_k)=\psi(x_k)\{\alpha'(x_k)+[\alpha(x_k)]^2\}$. For $x_k\leqq p$, $\alpha(x_k)\geqq 0$ and from (18) we see that $\alpha(x_k)<1/x_k$. Hence $\alpha'(x_k)+[\alpha(x_k)]^2$ is negative in this region. By continuity there is some $q>p$ such that this function, and hence $\psi''(x_k)$, remains negative for $x_k<q$. Thus ψ is concave in $[0, q]$.

Now let the maximum of $\sum\psi(x_k)$ occur at $\bar{x}_1\leqq\cdots\leqq\bar{x}_m$ and assume that $\bar{x}_m\geqq q$. Then necessarily $\bar{x}_1<\mu$ and hence, for sufficiently small $\epsilon>0$, $\psi(\bar{x}_1)<\psi(\bar{x}_1+\epsilon)$ and $\psi(\bar{x}_m)<\psi(\bar{x}_m-\epsilon)$. Clearly this is impossible; hence $\bar{x}_m<q$. Since $\bar{x}_1,\cdots,\bar{x}_m$ all fall in an interval where ψ is concave, a variation of the earlier argument shows that the maximum occurs at $\bar{x}_1=\cdots=\bar{x}_m=\mu$. Hence

$$(19)\qquad f_m(x_1,\cdots,x_m)\leqq\frac{1}{x_0+m\mu}m\psi(\mu)$$

from which (15) follows.

Reference

1. Edward O. Thorp, Repeated independent trials and a class of dice problems, this MONTHLY, 71 (1964) 778–781.

JACOBI ENDOMORPHISMS

B. M. PUTTASWAMAIAH, Carleton University, Ottawa

An endomorphism σ of an arbitrary group G will be called a *Jacobi endomorphism* if σ satisfies the Jacobi identity

$$((ab)^\sigma c)^\sigma((bc)^\sigma a)^\sigma((ca)^\sigma b)^\sigma=1$$

for all elements a, b and c in G, where a^σ is the image of a under σ. For a given group G, there always exists at least one Jacobi endomorphism, namely the zero or trivial endomorphism given by $a^0=1$ for all a in G. A group G which admits a nontrivial Jacobi endomorphism will be called a *Jacobi group*. In the following we show that G is a Jacobi group if and only if G is a semi-direct product of a normal subgroup and an abelian group of odd exponent. It is also shown that if G is an abelian group of odd exponent, then there exists a Jacobi automorphism.

If σ and ρ are two endomorphisms of G, we mean by the sum $\sigma+\rho$ the function which maps an element a of G onto $a^\sigma a^\rho$. The function $\sigma+\rho$ is an endo-

morphism of G if and only if G^σ and G^ρ are elementwise commutative [1]. If ϕ is any function of G into itself, we write $n|\phi$ to mean that n is a factor of ϕ.

A group G is said to be a *semi-direct product* of a proper normal subgroup H and a subgroup K if and only if the elements of K may be taken as the coset representatives of H in G.

LEMMA 1. *A group G has a Jacobi endomorphism σ if and only if*

$$n \mid 2\sigma^2 + \sigma \qquad \text{and} \qquad a^\sigma b^{\sigma^2} = b^{\sigma^2} a^\sigma \tag{1}$$

for all a, b in G, where n is the exponent of G^σ.

Proof. If σ is a Jacobi endomorphism of G, then by definition

$$((ab)^\sigma c)^\sigma((bc)^\sigma a)^\sigma((ca)^\sigma b)^\sigma = 1$$

for all elements a, b and c in G, which can also be written in the form

$$a^{\sigma^2} b^{\sigma^2} c^\sigma b^{\sigma^2} c^{\sigma^2} a^\sigma c^{\sigma^2} a^{\sigma^2} b^\sigma = 1. \tag{2}$$

In (2), set $a=c=1$, obtaining $b^{2\sigma^2+\sigma}=1$ for all b in G, which implies $n|2\sigma^2+\sigma$. On the other hand, letting $c=1$ in (2), we obtain $a^{\sigma^2}b^{\sigma^2}b^{\sigma^2}a^\sigma a^{\sigma^2}b^\sigma=1$ which in view of $n|2\sigma^2+\sigma$, can be written as

$$a^{\sigma^2} b^{-\sigma} a^{-\sigma^2} b^\sigma = 1 \qquad \text{or} \qquad a^{\sigma^2} b^\sigma = b^\sigma a^{\sigma^2}$$

for all a and b in G. This proves the first part of the Lemma.

Conversely let the conditions (1) hold. Then by repeated use of (1), we have

$$\begin{aligned}
((ab)^\sigma c)^\sigma((bc)^\sigma a)^\sigma((ca)^\sigma b)^\sigma &= a^{\sigma^2} b^{\sigma^2} c^\sigma b^{\sigma^2} c^{\sigma^2} a^\sigma c^{\sigma^2} a^{\sigma^2} b^\sigma \\
&= a^{\sigma^2} b^{\sigma^2} b^{\sigma^2} (c^\sigma c^{\sigma^2} c^{\sigma^2}) a^\sigma a^{\sigma^2} b^\sigma \\
&= a^{\sigma^2} b^{\sigma^2} b^{\sigma^2} (c^{\sigma^2} c^{\sigma^2} c^\sigma) a^\sigma b^\sigma a^{\sigma^2} \\
&= a^{\sigma^2} b^{\sigma^2} b^{\sigma^2} a^\sigma b^\sigma a^{\sigma^2} \\
&= a^{\sigma^2} a^\sigma (b^{\sigma^2} b^{\sigma^2} b^\sigma) a^{\sigma^2} \\
&= a^{\sigma^2} a^\sigma a^{\sigma^2} \\
&= 1
\end{aligned}$$

and thus σ is a Jacobi endomorphism of G.

COROLLARY. *If σ is a Jacobi automorphism of G, then G is abelian.*

THEOREM 1. *If σ is a Jacobi endomorphism of a group G then* (i) *G^σ is abelian,* (ii) *σ is an automorphism of G^σ,* (iii) *G is of odd exponent, and* (iv) *G is a semi-direct product of G^σ and the kernel of σ.*

Proof. Since σ is a Jacobi endomorphism of G, we have from (1) that $a^{\sigma^2}b^\sigma a^{-\sigma^2}=b^\sigma$ for all a and b in G. This implies that G^{σ^2} lies in the centre of G^σ. But the centre is abelian and therefore G^{σ^2} is abelian.

If x is an arbitrary element of G^σ, then $x=(y^{-1})^\sigma=y^{-\sigma}$ for some y in G. Since

$n\,|\,2\sigma^2+\sigma$, we can also write $x=y^{-\sigma}=y^{2\sigma^2}=(y^2)^{\sigma^2}$, which implies that x belongs to G^{σ^2}. Thus $G^\sigma=G^{\sigma^2}$ and so G^σ is abelian, and σ is an automorphism of G^σ.

Further the relation $n\,|\,2\sigma^2+\sigma$ implies that $G^{2\sigma^2}=G^{-\sigma}=(G^{-1})^\sigma=G^\sigma=G^{\sigma^2}$. Hence G^σ does not contain an element of order 2; and thus G^σ is an abelian group of odd exponent. Finally G^σ is isomorphic to the factor group $G/\ker(\sigma)$ implies that G is a semi-direct product of G^σ and $\ker(\sigma)$, thus proving the theorem.

In order to prove the converse, we require the following

LEMMA 2. *If G is an abelian group of odd exponent, then there exists a Jacobi automorphism.*

Proof. Let $2n+1$ be the exponent of G. Then $a^{2n+1}=1$ for all a in G. Define a mapping σ of G into itself by $a^\sigma=a^n$ for all a in G. Since G is an abelian group, the mapping σ is an endomorphism. The integers $2n$ and $2n+1$ being relatively prime, the relation $a^n=1$ for any a in G implies $a=a^{2n+1}=1$, so that the kernel of σ consists of the identity only. Thus σ is an automorphism. Since σ is nontrivial, it is an outer automorphism.

Also $2\sigma^2+\sigma=2n^2+n=n(2n+1)$, so that $2n+1\,|\,2\sigma^2+\sigma$. Therefore it follows from Lemma 1, that σ is a Jacobi automorphism and G is a Jacobi group.

Example. As an illustration consider the additive group Z_{2n+1} of integers Z modulo $2n+1$, where n is any positive integer. If we now define a mapping in Z_{2n+1} by $a^\sigma=ka$ where a is in Z_{2n+1} and k is in Z, then it is easily seen that σ is an automorphism if $k=n$. Since $2n^2+n=n(2n+1)$, the automorphism σ satisfies the hypothesis of Lemma 1. Hence Z_{2n+1} is a Jacobi group.

THEOREM 2. *If G is a semi-direct product of a normal subgroup H and an abelian group K of odd exponent, then G is a Jacobi group.*

Proof. Since H is a normal subgroup in G, the factor group G/H is well defined. Let ν be the natural homomorphism of G onto G/H given by $a^\nu=aH$. Also let i be the isomorphism of G/H onto K. Since G is a semi-direct product of H and K, every coset of H in G contains exactly one element of K. Then this isomorphism i can be described by $(aH)^i=a'$, where a' is the common element of K and aH. Let $2n+1$ be the exponent of K. The mapping ρ of K into itself defined by $a'^\rho=a'^n$ is a Jacobi automorphism (see the proof of Lemma 2).

If we set $\sigma=\nu i\rho$, then we shall show that σ has the desired property. We have

$$\begin{aligned}(ab)^\sigma &= (ab)^{\nu i\rho} = ((ab)^\nu)^{i\rho} = (abH)^{i\rho} = ((aH\cdot bH)^i)^\rho = ((aH)^i(bH)^i)^\rho = (a'b')^\rho \\ &= a'^\rho b'^\rho = ((aH)^i)^\rho((bH)^i)^\rho = (aH)^{i\rho}(bH)^{i\rho} = (a^\nu)^{i\rho}(b^\nu)^{i\rho} = a^{\nu i\rho}b^{\nu i\rho} = a^\sigma b^\sigma,\end{aligned}$$

so that σ is an endomorphism. Also

$$a^{2\sigma^2+\sigma} = a^{\sigma^2}a^{\sigma^2}a^\sigma = a'^{n^2}a'^{n^2}a'^n = a'^{2n^2+n} = a'^{n(2n+1)} = 1,$$

so that $2n+1\,|\,2\sigma^2+\sigma$. Finally the equations

$$a^\sigma b^{\sigma^2} = a'^n b'^{n^2} = b'^{n^2}a'^n \qquad \text{and} \qquad b^{\sigma^2}a^\sigma = b'^{n^2}a'^n$$

imply $a^{\sigma}b^{\sigma^2}=b^{\sigma^2}a^{\sigma}$ for all a and b in G. From Lemma **1**, it follows that σ is a Jacobi endomorphism and so G is a Jacobi group.

THEOREM 3. *If G is a Jacobi group under an inner automorphism σ of G, then G is a direct product of cyclic groups each of order* 3.

Proof. Let t be the element of G which causes the inner automorphism σ. Then the Jacobi identity takes the form

$$(t^{-2}abtct)(t^{-2}bctat)(t^{-2}catbt) = 1$$

for all a, b and c in G. In this identity if we let $a=t$, $b=c=1$, we obtain $t^3=1$, while $b=c=1$ gives $tat^{-1}=a^{-2}$. Finally $c=a^{-1}$ gives $ab=ba$ for all a and b in G. Thus G is an abelian group in which $a=a^{-2}$ for every a in G. Hence G is a direct product of cyclic groups each of order 3.

REMARKS. In (1), the equation $n\mid 2\sigma^2+\sigma$ has at most two solutions. Therefore a group G has at most two Jacobi endomorphisms, one of them being the zero endomorphism.

The Jacobi identity can be expressed in a slightly different form. An endomorphism ρ of G may be called Jacobi endomorphism if

$$((ab)^{\sigma}c)^{\rho}((bc)^{\sigma}a)^{\rho}((ca)^{\sigma}b)^{\rho} = 1$$

for all a, b and c in G, where σ is a fixed endomorphism of G. Then the conditions (1) take the form

$$n \mid 2\sigma\rho + \rho \qquad \text{and} \qquad a^{\sigma\rho}b^{\rho} = b^{\rho}a^{\sigma\rho}$$

for all a, b in G and where n is the exponent of G^{ρ}.

Finally, I would like to thank the referee for his suggestions.

Reference

1. H. J. Zassenhaus, The Theory of Groups, 2nd ed., Chelsea, New York, 1958.

SOME RELATIVES OF MINKOWSKI'S THEOREM FOR TWO-DIMENSIONAL LATTICES

JOSEPH HAMMER, University of Sydney

Introduction. Minkowski's Theorem states that any convex domain of area greater than four which is symmetric about a lattice point, contains at least two more lattice points.

M. E. Ehrhart [2] has generalized it in the following way: any convex domain of area greater than 4.5 of which the centre of gravity is a lattice point, encloses at least two more lattice points.

E. A. Bender [1] has proved that any convex domain whose area is greater than half its perimeter, contains a lattice point.

In this note, we shall show certain relations between these theorems and generalize them.

All theorems are understood to be valid in two-dimensional space. The first theorem is a relation between Minkowski's and Bender's theorems.

THEOREM 1. *Any convex domain whose perimeter is not greater than its area and symmetric about a lattice point, contains at least four more lattice points.*

To prove this we shall use the following lemma.

LEMMA 1. *Any convex domain whose perimeter P is not greater than its area, that is*

$$A \geqq P$$

can be divided into two convex domains, such that each area is greater than the corresponding half perimeter.

The statement of this lemma is a special case of the lemma in [**3**] if we take $n=2$. The only difference is that the condition "the area is greater than the perimeter" has been extended to "the perimeter is not greater than the area." The proof, with the extended condition, is identical with that in [**3**], since the domain is convex.

To prove the theorem, let us take a chord through the centre of the domain. This bisects the area, since it is central-symmetric. By Lemma 1, the area of each part is greater than the corresponding half perimeter. Applying Bender's Theorem we find that each part contains a lattice point. Now we have two cases: either these two points are not symmetrically located with respect to the centre, in which case it is clear that each lattice point has its symmetric counterpart; or, they are symmetrical points, in which case the chord through them goes through the centre and bisects the area. Since the area of each part is greater than the corresponding half perimeter, each part encloses a new lattice point.

Now, in Theorem 2, we will have a relation between Ehrhart's and Bender's theorems.

THEOREM 2. *Any convex domain whose perimeter is not greater than its area and whose centre of gravity is a lattice point contains at least two more lattice points.*

This theorem is a special case of the following theorem:

THEOREM 3. *Any convex domain whose perimeter is not greater than its area, encloses at least three lattice points.*

To prove this, we use the following lemma:

LEMMA 2. *Through any point in a convex domain, a chord can be drawn which bisects the area.*

Proof. Take a chord through the point. Suppose that this chord divides the domain into two unequal parts, A_1 and A_2; suppose that $A_1 - A_2 > 0$. As the chord does a half turn around the point, A_1 and A_2 vary continuously with the

angle of rotation, to exchange their magnitude, yielding $A_1 - A_2 < 0$. By Bolzano's Theorem on continuous functions, for one position at least, $A_1 = A_2$, what we wished to prove.

To prove the theorem, it is plain by Bender's Theorem that the domain encloses a lattice point. By Lemma 2, there exists an area bisector through this point. By Lemma 1, each area is greater than the corresponding half perimeter, thus each contains a lattice point, which differs, of course, from the lattice point lying on the bisector chord.

Now we prove a more general theorem.

THEOREM 4. *Any convex domain whose area is greater than or equal to 2^r-times its perimeter P, that is*

$$A \geqq 2^r P, \tag{1}$$

where r is a positive integer, contains at least $2^{r+2}-1$ lattice points.

We prove it by induction. Theorem 3 is the case when $r=0$. Supposing it is true till $r=n-1$, we show it for $r=n$.

Let us bisect the area by a chord which goes through a lattice point. From (1) for each half, we can write

$$A/2 = A' \geqq 2^{n-1}P.$$

But, since the domain is convex, it is clear that

$$A' \geqq 2^{n-1}P > 2^{n-1}P',$$

where P' is the perimeter of the half domain. Thus by the inductive hypothesis, each half contains at least $2^{n+1}-1$ lattice points, and the domain contains

$$2(2^{n+1} - 1) + 1 = 2^{n+2} - 1$$

lattice points.

References

1. Edward A. Bender, Area-perimeter relations for two-dimensional lattices, this MONTHLY, 69 (1962) 742–744.

2. M. Eugene Ehrhart, Comptes Rendus, Acad. Sci. Paris, 240 (1955) 483–485.

3. Joseph Hammer, On a general area-perimeter relation for two-dimensional lattices, this MONTHLY, 71 (1964) 534–535.

ON THE ZEROS OF SOLUTIONS TO BESSEL'S EQUATION

F. T. METCALF AND MILOŠ ZLÁMAL, University of Maryland

Consider Bessel's equation

$$\ddot{x}(t) + \frac{1}{t}\dot{x}(t) + \left(1 - \frac{\nu^2}{t^2}\right)x(t) = 0, \qquad 0 < t, \tag{B}$$

where ν is a real parameter. A classical theorem on Bessel functions states:

whenever $T_1 < t < T_2$. Integration between T_1 and T_2 then gives the results contained in the second statement of the theorem.

Suppose now that $0 < T \leqq T_1 < T_2$, where $T^2 > \nu^2 - \frac{1}{4}$. Then

$$-1 + \frac{\nu^2 - \frac{1}{4}}{T^2}\cos^2\theta(t) \quad \begin{matrix} < \dot{\theta}(t), & \nu^2 < \frac{1}{4}, \\ > \dot{\theta}(t), & \nu^2 > \frac{1}{4}, \end{matrix}$$

whenever $T_1 < t < T_2$. Since $T^2 > \nu^2 - \frac{1}{4}$, the left-hand side of each of the above inequalities is negative. Hence,

$$-\frac{\pi}{[1 - \{(\nu^2 - \frac{1}{4})/T^2\}]^{1/2}}$$
$$= \int_{+\pi/2}^{-\pi/2} \frac{d\phi}{1 - \{(\nu^2 - \frac{1}{4})/T^2\}\cos^2\phi} \quad \left\{ \begin{matrix} > -(T_2 - T_1), & \nu^2 < \frac{1}{4}, \\ < -(T_2 - T_1), & \nu^2 > \frac{1}{4}, \end{matrix} \right.$$

which yields the desired results.

The research of the first author was supported in part by the Air Force Office of Scientific Research under Grant AFOSR 400-64; while the research of the second author was supported by the National Science Foundation under Grant GP 2067.

References

1. G. N. Watson, Theory of Bessel Functions, Cambridge, 1962.
2. Frank Bowman, Introduction to Bessel Functions, Dover, New York, 1958.
3. Adam Czarnecki, A note on the zeros of Bessel functions, this MONTHLY, 71 (1964) 403–404.
4. E. A. Coddington and Norman Levinson, Theory of ordinary differential equations, McGraw-Hill, New York, 1955.
5. E. Kamke, Differentialgleichungen: Lösungsmethoden und Lösungen, Chelsea, New York, 1948.

A RODRIGUES' FORMULA FOR THE LAGUERRE POLYNOMIALS

L. R. BRAGG, Case Institute of Technology

In this note, we will be concerned with the development of a version of the Rodrigues' formula for the generalized Laguerre polynomials. This formula, which doesn't seem to appear in the literature, is given by

$$(1) \qquad L_j^{(\alpha)}(r^2) = \frac{(-1)^j}{2^{2j}j!}\{e^{r^2}\Delta_{2(\alpha+1)}^j e^{-r^2}\}, \qquad j = 0, 1, 2, \cdots.$$

Here, $\Delta_{2(\alpha+1)} \equiv D_r^2 + [(2\alpha+1)/r]D_r$, $D_r \equiv \partial/\partial r$, and this reduces to the usual "radial" Laplacian operator when $n = 2(\alpha+1)$ is a positive integer. The interest in the above formula is twofold: (a) it is a generalization of the form of the Rodrigues' formula for the Hermite polynomials ([3], p. 189) only involving the operator $\Delta_{2(\alpha+1)}$ and (b) it serves as a good starting point for easily deriving a number of basic formulas for these polynomials. For example, if we select $\alpha = -\frac{1}{2}$

and $\Delta_1 \equiv D_r^2$ in (1), we obtain the connection between the Hermite polynomials $H_{2j}(r)$ and the Laguerre polynomials $L_j^{-1/2}(r^2)$. On the other hand, the choices $\alpha=0$, $r^2=x_1^2+x_2^2$, and $\Delta_2 \equiv D_1^2+D_2^2$ in (1) along with an application of the binomial theorem leads to the relation

$$L_j^{(0)}(x_1^2+x_2^2) = \frac{1}{2^{2j}j!}\sum_{k=0}^{j}\binom{j}{k}H_{2(j-k)}(x_1)H_{2k}(x_2),$$

a result first announced by Feldheim [1].

Our development of (1) is based upon an elementary semi-group property for solutions of the heat equation and a generating function related to the source solution of this equation. A modification of this method can also be used to derive the Rodrigues' formula for the Hermite polynomials. The result (1) can also be derived directly from the basic recurrence relations for the Laguerre polynomials ([3] p. 202), but we omit the use of these here.

Let us consider the initial-value generalized heat problem

$$(2)\qquad \begin{cases} \text{(a)}\ U_t(r,t) = \Delta_\mu U(r,t), & \mu = 2(\alpha+1) > 0 \\ \text{(b)}\ U(r,0) = \phi(r). \end{cases}$$

It is not assumed that μ is an integer. Nevertheless, we can attach to this problem a symbolic solution operator $\exp(t\Delta_\mu)$ (see Chapter 8, [2] when $\mu=1$) such that the analytic solution of (2) is given by

$$(3)\qquad U(r,t) = \exp(t\Delta_\mu)\cdot\phi(r),$$

provided that $\phi(r)$ satisfies suitable growth conditions (the Goursat conditions)· It is frequently necessary to restrict the time variable t in order that (3) be meaningful. These restrictions then define a domain $D(U)$ of validity of (3). It will be assumed here that if (r, t_1), (r, t_2), and (r, t_1+t_2) are any three points of $D(U)$, then we have the *semi-group property*

$$(4)\qquad \exp\{t_1\Delta_\mu\}\cdot U(r,t_2) = U(r, t_1+t_2).$$

Further, let $S(r, t)=(4\pi t)^{-\mu/2}e^{-r^2/4t}$, the so-called *source solution* of (2a). If $U(r, t)$ is a solution of (2a) in $D(U)$ and $(r/t, -1/t)\in D(U)$, then we define the *Appell transform* $\tilde{U}(r, t)$ of $U(r, t)$ by

$$(5)\qquad \tilde{U}(r,t) = S(r,t)U(r/t, -1/t).$$

Then $\tilde{U}(r, t)$ also satisfies (2a).

Let us now observe that if $\phi(r)=e^{ar^2}$ in (2b), then the corresponding solution of (2) is

$$U_a(r,t) = (1-4at)^{-\mu/2}\exp\{ar^2/(1-4at)\},$$

provided $1-4at>0$. If we select $T=4at$ and $X=-r^2/4t$ in this, its right-member reduces to the generating function for the Laguerre polynomials

$L_j^{(\mu/2-1)}(X)$ ([3] p. 202) and we have

$$U_a(r, t) = \sum_{j=0}^{\infty} T^j L_j^{'(\alpha)}(X) = \sum_{j=0}^{\infty} (4t)^j L_j^{(\alpha)}(-r^2/4t)a^j. \tag{6}$$

Upon forming the Appell transform of both sides of (6), we obtain

$$\begin{aligned} S(r, t + 4a) &= \tilde{U}_a(r, t) \\ &= S(r, t) \sum_{j=0}^{\infty} (-4/t)^j L_j^{(\alpha)}(r^2/4t)a^j. \end{aligned} \tag{7}$$

But since $S(r, t+4a) = \exp(4a\Delta_\mu) \cdot S(r, t)$ by (3), it follows, if we choose $t = \frac{1}{4}$ in (7), that

$$\exp(4a\Delta_\mu) \cdot S(r, \tfrac{1}{4}) = S(r, \tfrac{1}{4}) \sum_{j=0}^{\infty} (-16)^j L_j^{(\alpha)}(r^2)a^j. \tag{8}$$

Finally, if we evaluate the jth derivative of this with respect to a at $a = 0$, we obtain

$$(4\Delta_\mu)^j \cdot S(r, \tfrac{1}{4}) = (-1)^j j! 4^{2j} S(r, \tfrac{1}{4}) L_j^{(\alpha)}(r^2)$$

and this is just (1).

The research in this paper was supported by the National Aeronautics and Space Administration Grant No. NsG-544.

References

1. E. Feldheim, On Laguerre and Hermite Polynomials, Quart. J. Math., 11 (1940) 18–29.

2. I. I. Hirschman and D. V. Widder, The Convolution Transform, Princeton University Press, 1955.

3. E. Rainville, Special Functions, Macmillan, New York, 1960.

Editorial Note. In a private communication to Mr. V. K. Rohatgi, Professor H. W. Gould has pointed out in connection with Rohatgi's note, "Some Combinatorial Identities Involving Lattice Points," this MONTHLY, 73(1966), 507–508, that more general results than formula (3) in this note are given in the paper of Gould referred to in the note. Also he remarked that formula (3) holds for all real or complex numbers.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06467.

UNIQUENESS THEOREMS FOR A DIFFERENTIAL EQUATION

J. B. GARNER, Louisiana Polytechnic Institute

In a recent paper [2] the existence of at least one C^2 solution of

$$y'' + yF(y^2, x) = 0, \quad y(a) = y'(b) = 0, \quad y(x) > 0, \quad x \in (a, b]$$

was established by imposing certain conditions on $F(t, x)$. The uniqueness of this boundary value problem was established in [1]. The purpose of this note is to present to the student of differential equations elementary uniqueness theorems for similar systems.

We first establish uniqueness for a solution of a one-point boundary value problem.

THEOREM 1. *Let*
(A) *$F(t, x)$ be continuous in (t, x) for $x \in [a, b]$, $t \in [0, c]$,*
(B) *$F(t, x) > 0$ for $x \in [a, b]$, $t \in [0, c]$, and*
(C) *$t_2F(t_2, x) > t_1F(t_1, x)$ for $x \in [a, b]$ and $0 \leqq t_1 < t_2 < c$.*

Then there exists at most one C^2 solution of

$$y'' + yF(y, x) = 0 \tag{1}$$

satisfying

$$y(a) = \beta_1, \quad y'(a) = \beta_2, \quad y(x) > 0, \quad x \in (a, b], \tag{2}$$

where β_1, β_2 are nonnegative real numbers, both not zero, and

$$\beta_1 + (b - a)\beta_2 \leqq c. \tag{3}$$

Proof. Let $y_1(x)$ and $y_2(x)$ be two distinct solutions of (1), (2). Then, since these are C^2 solutions, we may assume

$$y_1(x) \equiv y_2(x), \quad x \in [a, t_1] \quad \text{and} \quad y_2(x) > y_1(x), \quad x \in (t_1, t_2) \tag{4}$$

for some t_1, t_2 of $[a, b]$. If $a = t_1$, $y_2'(t_1) = y_1'(t_1)$ by (2). If $a < t_1$ this equality follows from the continuity of $y_2' - y_1'$ and the first condition of (4). Hence,

$$y_2'(t_1) = y_1'(t_1), \quad y_2'(x) > y_1'(x), \quad x \in (t_1, t_3) \tag{5}$$

for some t_3 with $t_3 \leqq t_2$.

From (1) we have

$$y'(x) = y'(t_1) - \int_{t_1}^{x} y(u)F(y, u)du. \tag{6}$$

On writing (6) for y_1 and y_2 and making use of (5), we find

$$y_2'(x) - y_1'(x) = \int_{t_1}^{x} [y_1(u)F(y_1, u) - y_2(u)F(y_2, u)]du. \tag{7}$$

From (B) and (1) we have $yy''<0$ for $y\neq 0$. Hence all solution curves of (1) are concave toward the horizontal axis. Thus, if $y(x)$ does not change sign in $[a, b]$, the curve $y=y(x)$ must lie between the x-axis and the tangent to the curve at any point of the interval. We now have, by (3),

$$y_i(x) \leqq \beta_1 + (x - a)\beta_2 \leqq c \qquad x \in [a, b],$$

and can use (C) and (4) to conclude that the right member of (7) is negative for $x=x_1$, $t_1<x_1<t_3$. But, by (5), the left member is positive.

We now have a contradiction to the assumption $y_1(x)\not\equiv y_2(x)$ and the theorem follows.

Proved in a similar manner is

THEOREM 2. *Let* $F(t, x)$ *satisfy* (A), (B), (C). *Then there exists at most one* C^2 *solution of* (1) *satisfying*

$$y(b) = \beta_1,\ y'(b) = \beta_2, \qquad y(x) > 0,\ x \in (a, b)$$

where $0<\beta_1\leqq c$.

We now consider uniqueness for solutions of some two-point boundary value problems.

THEOREM 3. *Let* $F(t, x)$ *be continuous in* (t, x) *for* $x\in[a, b]$ *and* $0\leqq t<\infty$ *and satisfy*

$$t_2F(t_2, x) < t_1F(t_1, x) \qquad x \in [a, b],\ 0 < t_1 < t_2 < \infty. \tag{C$'$}$$

Then there exists at most one C^2 *solution of* (1) *satisfying*

$$y(a) = \beta_1,\ y'(b) = \beta_2, \qquad y(x) > 0,\ x \in (a, b], \tag{8}$$

where β_1, $\beta_2\geqq 0$ *and both not zero.*

Proof. Assume $y_1(x)$ and $y_2(x)$ satisfy (1), (8) and are distinct. Then there exists at least one subinterval (t_1, t_2) of (a, b) in which $y_1(x)\neq y_2(x)$. Also, by (8), we may choose t_1 and t_2 such that

$$y_1'(t_1) \leqq y_2'(t_1),\ y_1'(t_2) \geqq y_2'(t_2), \qquad y_2(x) > y_1(x),\ x \in (t_1, t_2). \tag{9}$$

From (6) we obtain

$$y_2'(t_2) - y_1'(t_2) = y_2'(t_1) - y_1'(t_1) + \int_{t_1}^{t_2} [y_1(x)F(y_1, x) - y_2(x)F(y_2, x)]dx. \tag{10}$$

We use (C′) and (9) to conclude that the right member of (10) is positive. But (9) implies that the left member is nonpositive. Hence, we must have $y_1(x) \equiv y_2(x)$.

Proved in a similar manner is

THEOREM 4. *Let $F(t, x)$ satisfy the hypotheses of Theorem 3. Then there exists at most one C^2 solution of (1) satisfying either*

(i) $y'(a)=\beta_1$, $y(b)=\beta_2$,
(ii) $y(a)=\beta_1$, $y(b)=\beta_2$, *or*
(iii) $y'(a)=\beta_1$, $y'(b)=\beta_2$,

where, in each problem, $y(x)>0$ on (a, b) and $\beta_1, \beta_2 \geqq 0$ with both not zero.

This work was done during the course of research sponsored by NSF grant GP-1337.

References

1. R. M. Moroney, Note on a Theorem of Nehari, Proc. Amer. Math. Soc., 13 (1962) 407–410.
2. Zeev Nehari, On a Class of Nonlinear Second-Order Differential Equations, Trans. Amer. Math. Soc., 95 (1960) 101–123.

THE REAL ROOTS OF THE EQUATION $(x-1)^k = x^{k-1}$

D. A. SMITH, Duke University

1. Introduction. There do not seem to be many examples of applications of the basic theorems of the calculus, such as the mean and intermediate value theorems, which are both accessible to the first-year student and illustrative of the power of these theorems. Some of the Sturm theorems for differential equations (e.g. if f and g are real linearly independent solutions of a second order linear homogeneous differential equation, then the zeros of f and g occur alternately) have simple proofs based on the mean value theorem, but there is the problem of explaining exactly what the statements mean. On the other hand, the error estimate for Simpson's rule is simply stated, but the student has difficulty following the details of the proof.

This note presents a simply-stated, nontrivial problem and some answers which are interesting in their own right and also readily accessible to the student of calculus. In addition to the mean and intermediate value theorems, we use the properties of monotone functions, approximation by early terms of a series, and other elementary ideas. The notation of number theory is used in places as a shorthand device, but this can be avoided.

2. The problem. For a given positive integer k, how large must x be so that $(x-1)^k > x^{k-1}$? Since we are asking for the largest root of the equation

$$(x-1)^k = x^{k-1}, \tag{1}$$

we may as well ask for all the roots. One quickly finds that there are two roots for

CLOSURE CONTINUITY

D. R. Andrew, University of Southwestern Louisiana, and Elaine Kirley Whittlesy, Iowa State Teachers College

In this note we compare a type of function, which we shall call a "closure continuous" function, with a continuous function. We show that in one sense continuity is a stronger property than closure continuity (Theorem 1) while in another sense it is not (Example 2). Conditions are obtained under which closure continuity is equivalent to continuity.

The notation "$f\colon X \to X^*$" will mean that f is a single-valued function from the topological space X to the topological space X^*.

Definition 1. *A function $f\colon X \to X^*$ is closure continuous at $x \in X$ if and only if for every open set G^* in X^* such that $f(x) \in G^*$, there is an open set G in X such that $x \in G$ and $f(cG) \subseteq c^*G^*$ (c and c^* denote the respective closure operators).*

Definition 2. *The function $f\colon X \to X^*$ is closure continuous (on X) if and only if f is closure continuous at every $x \in X$.*

Theorem 1. *If $f\colon X \to X^*$ is continuous on X, then $f\colon X \to X^*$ is closure continuous on X.*

Proof. Suppose that $f\colon X \to X^*$ is continuous on X. Let $x \in X$ and let G^* be any open set for which $f(x) \in G^*$. Since f is continuous, there is an open set G such that $x \in G$ and $f(G) \subseteq G^*$. Hence $c^*f(G) \subseteq c^*G^*$. By the continuity of f, $f(cG) \subseteq c^*f(G)$. Hence $f(cG) \subseteq c^*G^*$ and f is closure continuous on X.

Example 1. $f\colon X \to X^*$ may be closure continuous on X while discontinuous at every $x \in X$. Let $X(X^*)$ be the unit interval with topology consisting of the empty set together with all sets whose complements are finite (countable). Let $f\colon X \to X^*$ be the identity transformation. Then f is closure continuous since for every nonempty open set G^* in X^*, $c^*G^* = X^*$. It is clear that for every $x \in X$, f is not continuous at x.

Theorem 2. *Let X^* be a regular space. Then $f\colon X \to X^*$ is closure continuous on X if and only if $f\colon X \to X^*$ is continuous on X.*

Proof. Suppose that $f\colon X \to X^*$ is closure continuous on X. Let $x \in X$ and let G^* be any open set in X^* such that $f(x) \in G^*$. Since X^* is regular there is an open set G_1^* in X^* such that $f(x) \in G_1^*$ and $c^*G_1^* \subseteq G^*$. There is an open set G in X such that $x \in G$ and $f(cG) \subseteq c^*G_1^*$. It follows that $f(G) \subseteq G^*$. Hence f is continuous on X.

If $f\colon X \to X^*$ is continuous on X, then it is closure continuous on X by Theorem 1.

Corollary 1. *If X^* is regular and $f\colon X \to X^*$ is closure continuous at $x \in X$, then f is continuous at x.*

EXAMPLE 2. Given a point $x \in X$, a function $f\colon X \to X^*$ may be continuous at x and not closure continuous at x. Let $X=\{a, b, c\}$ with topology consisting of $\{a\}$, $\{a, b\}$, X, and $\varnothing$. Let $X^*=\{a, b, c\}$ with topology consisting of $\{a, b\}$, $\{c\}$, X^*, and $\varnothing$. Define $f\colon X \to X^*$ to be the identity mapping. Clearly f is continuous at b. That f is not closure continuous at b follows from a consideration of $G^*=\{a, b\}$ in X^*.

It follows from Theorem 1 that the function in the previous example cannot be continuous at both a and c. In fact we have the following

THEOREM 3. *If $f\colon X \to X^*$ is continuous at $x \in X$ but not closure continuous at x, then every open set containing x has a limit point which is a point of discontinuity of f.*

Proof. Let x be a fixed point of X and suppose that $f\colon X \to X^*$ is continuous at x but not closure continuous at x.

There exists an open set G_1^* such that $f(x) \in G_1^*$ and for every open set G for which $x \in G$, $f(cG) \not\subseteq c^*G_1^*$. Since f is continuous at x, there is an open set G_2 such that $x \in G_2$ and $f(G_2) \subseteq G_1^*$.

Let G_1 be any open set such that $x \in G_1$ and define $G_3=G_1 \cap G_2$. Then $x \in G_3$, $f(G_3) \subseteq G_1^*$, and $f(cG_3) \not\subseteq c^*G_1^*$. Hence there exists $y \in X$ such that y is a limit point of G_3 and $f(y) \notin c^*G_1^*$. Since y is a limit point of G_3, y is a limit point of G_1.

We show now that f is not continuous at y. The set $G_2^*=X^*-c^*G_1^*$ is open in X^* and $f(y) \in G_2^*$. Let G be an arbitrary open set in X such that $y \in G$. There is a point $z \in G_3$ such that $z \in G$. Hence $f(z) \in G_1^*$ and this implies $f(z) \notin G_2^*$ and therefore $f(G) \not\subseteq G_2^*$. This completes the proof.

THEOREM 4. *Let X be a regular space. If $f\colon X \to X^*$ is continuous at $x \in X$, then $f\colon X \to X^*$ is closure continuous at x.*

Proof. Suppose $f\colon X \to X^*$ is continuous at x and that G^* is any open set containing $f(x)$. There is an open set G such that $x \in G$ and $f(G) \subseteq G^*$. Since X is regular, there is an open set G_1 such that $x \in G_1$ and $cG_1 \subseteq G$. Hence $f(cG_1) \subseteq c^*G^*$, and f is closure continuous at x.

The following example shows that the converse of Theorem 4 does not hold.

EXAMPLE 3. Let $f\colon X \to X^*$ be as in Example 2. X^* is regular, and the function $f^{-1}\colon X^* \to X$ is closure continuous on X^* since for every nonempty open set G in X, $cG=X$. But f^{-1} is not continuous at the point $a \in X^*$.

We conclude with the following theorem which is an immediate consequence of Corollary 1 and Theorem 4.

THEOREM 5. *Let X and X^* be regular spaces and $x \in X$. Then $f\colon X \to X^*$ is closure continuous at x if and only if $f\colon X \to X^*$ is continuous at x.*

Reference

1. William J. Pervin, Foundations of General Topology, Academic Press, New York, 1964.

ON THE CONVERGENCE OF THE BINOMIAL SERIES

AUGHTUM HOWARD, Eastern Kentucky University

The binomial series, $1+\sum_{n=1}^{\infty}\binom{m}{n}x^n$, $m\neq 0, 1, 2, \cdots$, is the Maclaurin series representing the function $(1+x)^m$. It is easily shown to be convergent for $|x|<1$. The elementary calculus, however, usually gives only limited consideration to the question of convergence or divergence at $x=\pm 1$. Here the series becomes a function of m alone and is well known to be convergent for $m>-1$ at $x=1$ and for $m>0$ at $x=-1$.

This note gives an elementary proof of the convergence of the series for $-1<m<1$ at $x=1$ and for $0<m<1$ at $x=-1$. The method is a generalization of that used in [1] for $m=-1/2$, $x=1$. We shall need the

LEMMA. $[(a-1)/a]^p<a/(a+p)$, $a>1$, *p any integer* $\geqq 1$.

Proof. Using only the first two terms of the binomial expansion, we have

$$\left(\frac{a}{a-1}\right)^p=\left(1+\frac{1}{a-1}\right)^p\geqq 1+\frac{p}{a-1}>1+\frac{p}{a}=\frac{a+p}{a}>0.$$

The lemma follows by inverting this inequality.

Now, for $0<m<1$, there exists an integer $p\geqq 2$ such that $1/p<m$. For this p, we have

$$\left|\binom{m}{n}\right|=\frac{m(1-m)\cdots(n-1-m)}{n!}<1\left(1-\frac{1}{p}\right)\cdots\left(n-1-\frac{1}{p}\right)\frac{1}{n!}$$

$$=\frac{p-1}{p}\cdot\frac{2p-1}{2p}\cdots\frac{(n-1)p-1}{(n-1)p}\cdot\frac{1}{n}=a_n\cdot\frac{1}{n}\cdot$$

By the lemma,

$$(a_n)^{p-1}<\frac{p}{2p-1}\cdot\frac{2p}{3p-1}\cdots\frac{(n-1)p}{np-1}$$

$$=\frac{p}{p-1}\cdot\frac{2p}{2p-1}\cdots\frac{(n-1)p}{(n-1)p-1}\cdot\frac{p-1}{np-1}\leqq\frac{1}{a_n}\cdot\frac{1}{n}\cdot$$

Hence, $a_n<1/n^{1/p}$ and

$$\left|\binom{m}{n}\right|<1/n^{1+(1/p)}.$$

Since $1/n^{1+(1/p)}$ is the nth term of the well-known p-series, and $1+(1/p)>1$, we see that

$$1+\sum_{n=1}^{\infty}\binom{m}{n}x^n$$

is convergent at $|x|=1$ for $0<m<1$.

For $x=1$, $-1<m<0$, we give the usual test for convergence of an alternating series. For this range of m, it is easily shown that

$$\left|\binom{m}{n+1}\right| < \left|\binom{m}{n}\right|.$$

The series will then be convergent if

$$\lim_{n\to\infty}\left|\binom{m}{n}\right| = 0.$$

Let $-m=1-k$, $0<k<1$. Then

$$\left|\binom{m}{n}\right| = \frac{(1-k)(2-k)\cdots(n-1-k)}{(n-1)!}\cdot\frac{n-k}{n}.$$

By an argument similar to that used above for $0<m<1$, there is an integer $p\geqq 2$ for which the first factor on the right $<1/n^{1/p}$. Hence

$$0 \leqq \lim_{n\to\infty}\left|\binom{m}{n}\right| \leqq \lim_{n\to\infty}\frac{1}{n^{1/p}}\cdot\lim_{n\to\infty}\frac{n-k}{n} = 0\cdot 1 = 0.$$

Reference

1. G. E. F. Sherwood and A. E. Taylor, Calculus, Prentice-Hall, Englewood Cliffs, N. J., 1954, chapter XVI.

REAL PERIODIC FUNCTIONS

R. H. Cox and L. C. Kurtz, University of Kentucky

The purpose of this note is to give an elementary application of group theory to the theory of real periodic functions.

We will denote by $\mathfrak{R}$ the real number field, and by $\mathfrak{F}$ the set of functions f: $\mathfrak{R}\to\mathfrak{R}$. For each $f\in\mathfrak{F}$, let $G_f=\{p\,|\,p\in\mathfrak{R},\ f(t+p)=f(t)$ for all $t\in\mathfrak{R}\}$. It is clear that $0\in G_f$ for each $f\in\mathfrak{F}$. If $p\in G_f$ and $q\in G_f$, then

$$f(t+p-q) = f(t-q) = f(t-q+q) = f(t)$$

so that $p-q\in G_f$. Thus G_f is an additive subgroup of $\mathfrak{R}$. We will call G_f the group of periods of f.

Definition 1. *A function $f\in\mathfrak{F}$ is said to be periodic if G_f contains more than one element.*

It is a well known (and easily proved) theorem that every nontrivial additive subgroup G of $\mathfrak{R}$ is one of the following types:

(A) There exists a least positive element $p_0\in G$ and $G=[p_0]$, the cyclic group generated by p_0.

(B) The subgroup G is dense in $\mathcal{R}$ (in the usual topology).

DEFINITION 2. *A function $f \in \mathfrak{F}$ is said to have principal period p_0 if f is periodic, G_f is of type A, and p_0 is the least positive element of G_f.*

THEOREM 1. *If $f \in \mathfrak{F}$ is periodic, and if there exists a point $t_0 \in \mathcal{R}$ such that f is continuous at t_0, then either f is a constant function or f has a principal period.*

Proof. Suppose that f has no principal period, so that G_f is dense in $\mathcal{R}$. If $\epsilon > 0$, there exists $\delta > 0$ such that $|f(t_1) - f(t_0)| < \epsilon$ when $|t_1 - t_0| < \delta$. If $t \in \mathcal{R}$, we may express $t = t_0 + t' + t''$, where $t' \in G_f$ and $|t''| < \delta$, since G_f is dense in $\mathcal{R}$. Then

$$|f(t) - f(t_0)| = |f(t_0 + t' + t'') - f(t_0)| = |f(t_0 + t'') - f(t_0)| < \epsilon.$$

This shows that $f(t) = f(t_0)$, and f is a constant function.

The "converse" of Theorem 1 is false. An example is the function f given by $f(t) = 0$ (t rational) and $f(t) = 2 + \cos 2\pi t$ (t irrational). This function has principal period 1 but is everywhere discontinuous.

We now turn our attention to the functions of type B. Suppose that G is a subgroup of $\mathcal{R}$, and consider the factor group $\mathcal{R}/G$. If $\hat{f}$ is a function from $\mathcal{R}/G$ to $\mathcal{R}$, we may use $\hat{f}$ to induce a function $f \in \mathfrak{F}$ whose period group is G by defining $f(t) = \hat{f}(t + G)$. In particular, if G is the subgroup R of rationals, it is clear that the cardinality of $\mathcal{R}/R$ is $\aleph$. In this case we may choose f as a 1-1 mapping of $\mathcal{R}/R$ onto $\mathcal{R}$, and the resulting function f is of type B and assumes uncountably many values. This construction will easily yield a periodic function whose graph is dense in the plane. As a final remark, we observe that there exist uncountable proper subgroups of $\mathcal{R}$, and hence there exist non-constant periodic functions with uncountably many distinct periods.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and
JOHN A. BROWN, University of Delaware

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

A DILEMMA IN DEFINITION

C. P. NICHOLAS, United States Military Academy

An unfortunate tendency of today's textbooks on calculus is to place the student in a quandary over the term "function." The definition of this term which had long commanded professional and academic acceptance is being supplanted in recent texts by substitute versions which—although heralded as improvements in precision—fail notably on grounds of consistency. The re-

sult is an ambiguity which even the most careful mathematical writers are unable to avoid. There is danger that the new definitions are producing an irreconcilable conflict with a meaning long established in the literature of science and engineering. If this tendency continues, a student's first course in calculus seems destined to become a bewildering encounter with inconsistent language and overly elaborate symbolism.

In order to discuss the problem in specific terms, it will be helpful to state the conflicting definitions and assign corresponding marginal references. The definition of function that was widely accepted until about ten years ago appeared in substantially the following terms:

(A) *A function is a variable so related to another variable that to each value of the latter there corresponds uniquely a value of the former.*

In contrast, the definition tending to prevail in the most recently published calculus texts would be as follows:

(B) *A function is a set of ordered pairs, no two of which have the same first component.*

Still another view regards a function as a rule of correspondence. There are many ways of stating this, but the following paraphrase is sufficiently explicit for later reference:

(C) *Given a set of ordered pairs such that no two have the same first component, the rule of correspondence which associates each second component with the corresponding first is called a function.*

Certain variants of these definitions could be added to the list, but it will suffice to focus attention on (A), (B) and (C).

Now, under all three definitions the symbol $f(x)$ is customarily used to denote what is called the *value* of the function. Further, if we let (x, y) denote an ordered pair in the sense of (B) and (C), then under all three definitions it is customary also to consider y as the value of the function. Thus y and $f(x)$ are given the same meaning, so that in all three systems it is acceptable to write $y \equiv f(x)$. Regardless of which definition is chosen, therefore, the symbolism prevalent in science and engineering may be used whenever one's only concern is the function-value corresponding to a selected argument-value. Accordingly, if the only purpose of mathematical education were to prepare the student for stereotyped problems, one might say that the fundamental conflict of language between (A), (B) and (C) is of small consequence. But to take this attitude would be to revive the cookbook approach that the recent reforms have been intended to remedy.

If each of the definitions (A), (B) and (C) be reduced to the essential single word, we find that (A) calls a function a variable while (B) calls it a set, and (C) calls it a rule. These three meanings are not logically equivalent; and beyond that, (B) and (C) are at variance with the meaning of function most widely accepted in applications of calculus to science and engineering. By this I do not mean to belittle the concepts defined in (B) and (C); I am fully aware that these are important objects of interest, and that each deserves a standard term in

mathematical literature. My only point is that neither of these should be referred to as a function, since this terminology conflicts not only with a previously accepted mathematical definition but also with a fundamental meaning established in the English language by prolonged good usage.

As to what terms I would propose instead for (B) and (C), well suited words are available, but there is little point in discussing these unless teachers first agree that the overthrow of definition (A) is neither necessary nor desirable. Definition (A) is fully adapted to a precise treatment of calculus, and the usage implied by that definition is so deeply imbedded in the literature that it is not likely to vanish in our lifetime. Agreement on these points would make it clear that the term "function" cannot reasonably be applied to (B) or (C), if calculus is to be kept free of ambiguity. The problem would then reduce primarily to that of choosing an acceptable name for the concept denoted by the symbol f when used separately from $f(x)$. A reasonable solution to this problem was suggested as far back as 1953 by J. Barkley Rosser, whom I shall now quote. In his admirable *Logic for Mathematicians*, (McGraw-Hill, 1953), Rosser offers on page 309 the following suggestion:

> Actually, a perfectly good name for the notion f is available, namely "transformation." In algebraic geometry, a careful distinction is usually made between a "transformation" f and a "general value" $f(x)$ of the transformation. Thus, one way out of our impasse would be always to refer to f as a transformation and to reserve the term "function" to refer to $f(x)$.

This suggestion in Rosser's book occurs in the course of an illuminating discussion of the dilemma which is the subject of the present article. Pointing out that f and $f(x)$ are different and that it is confusing to apply the term "function" to both, he continues as follows:

> One should definitely decide to call one by the name "function" and then devise a new name for the other. The present trend in higher mathematics is to reserve the name "function" for f, but even those who advocate this are usually inconsistent in their use of the word "function."

Rosser proceeds from this point with a description of certain difficulties raised by calling $f(x)$ a function, but after close study I am persuaded that the problem is a minor one whose solution requires no more formidable remedy than careful English. Moreover, Rosser shows convincingly that symbolic logic provides adequate machinery for considering $f(x)$ a function and treating it in that sense with complete precision.

After describing various alternatives, Rosser reaches the point where a decision is necessary in order to proceed with the treatment of functions in his own pages. For this purpose he accepts the emerging practice of calling f the function, and in order to preserve consistency he is careful thereafter to refer to $f(x)$ as the function value. He points out that the procedure thus adopted is not highly satisfactory, but seems most nearly in accord with the present trend of mathematical thought.

In order to illustrate the growing dilemma I shall cite a few examples of more recent date and at the same time try to clarify the nature of the problem. An

important source of trouble is doubt in the minds of certain writers that the concept of a variable can be made precise. A person who does not know what a variable is would, of course, not understand definition (A). Some authors try to avoid the variable concept altogether, a case in point being the writer of a prominent introductory treatment of functions published in 1960. At one stage he remarks, "For the purposes of this book it is unnecessary to attempt to define 'a variable,' and we shall neither define nor use this word." Later on, however, he freely uses the verbs "increase" and "decrease," and he does not hesitate to state that certain functions are increasing or decreasing. Thus he has used the *idea* of a variable while avoiding the word. This in itself might be regarded as not serious, except for the fact that he had previously defined a function as a set of ordered pairs. Thus, he seems to be saying that a set of ordered pairs is increasing or decreasing.

I do not believe, of course, that he meant this, nor would an experienced mathematician insist on this literal interpretation. His book is intended for readers who have completed at least a course in calculus. Nevertheless, even at the level for which he was writing, the need to discuss functional behavior in intelligible language forced him to revert to the forms of speech which treat a function as a variable. Moreover, it is clear that he was aware of the language problem, for shortly afterward he comments that the function whose value at x is $\log x$ would usually be called "the function $\log x$." Proceeding then to the function whose value at x is $\log(\sin x)$, he states:

> Here we are forced into a certain amount of awkwardness if we are to avoid the possibly misleading "the function $\log (\sin x)$." The same problem arises when we want to talk about functions that are so simple that they have no generally accepted names.

In view of the unconcealed conflict of language in this treatment, a student who reads it would be left in serious doubt as to whether such entities as x^2, $\sin x$, $\log(\sin x)$, and so on, may properly be referred to as functions.

Several modern authors have been frank to acknowledge the dilemma, and after defining a function in accordance with the new vogue, they revert openly to the variable concept in order to provide intelligible discussion. Thus the preface of a certain text on calculus published in 1964 includes the following confession:

> The notation $f(x)$ is defined with complete propriety as the image of the function f for a given pre-image x: that is, $y=f(x) \Leftrightarrow (x, y) \in f$. But to avoid too great a clash with tradition and with the bulk of mathematical literature, we later slip into the (bad?) habit of calling $f(x)$ a function.

The text then proceeds, commendably in my opinion, to refer to such entities as x^2, x^3, $\sin x$, and so on, as functions.

There are other authors who have gone to great lengths to avoid such overt inconsistency, but the inevitable consequence is excessive symbolism and circumlocution. This is well illustrated by a textbook on calculus written by two prominent authors and published in 1962. These writers define a function to be a set of ordered pairs no two of which have the same first component, and as a

typical symbol they employ the letter F. Their preface introduces the following conventions for distinguishing between F and $F(x)$:

> If F denotes a function, and if $(x, y)\in F$, we call y the *correspondent* of x under F, and we denote this correspondent by $F(x)$. We most commonly specify a function by writing
>
> $$F = \{(x, y) \mid y = F(x)\},$$
>
> and if $(a, F(a))\in F$, we call $F(a)$ the value of $F(x)$ at a.

This device enables its authors to avoid the logical inconsistency of first saying that a function F is a set, and then saying that $F(a)$ denotes a value of the function (literally, a "value of the set F"). Instead, they say that $F(a)$ is a value of the *correspondent*, thus ascribing to the latter concept the intrinsic character of a variable.

While I admire the sincerity with which these authors have insisted on consistency, I feel nevertheless that they have demanded of the student too high a price. The student who uses this book will be dealing, not only with functions, but also with correspondents; and not only with derivatives of functions, but also with derivatives of correspondents; and not only with antiderivatives of functions, but also with antiderivatives of correspondents. As for notation, in cases where educated men have long felt that the display

$$\sin x$$

denotes a function, the student of this 1962 text will be required instead to denote the function by the display

$$\sin = \{(x, y) \mid y = \sin x\}.$$

Since a function is a set of ordered pairs, this student may not say that a function is capable of increasing or decreasing or having an extreme value.

Now, without discussing the mathematical merits of these dual conventions, I must view them with regret from the standpoint of good teaching. Calculus can be made simpler than this without impairing either its rigor or its precision. Moreover, a student who masters these conventions may find himself in a strange and conflicting world of applications. His physics teacher will probably persist in calling $\sin x$ a function, and may also insist that the distance traveled by a falling stone is a function of the time elapsed since the stone was released. The latter observation may cause the student to ponder how a distance can possibly be a set of ordered pairs. Such conflicts between mathematical fiat and the established terms of science and engineering are not likely to advance the cause of mathematical education. Rather, it seems to me, calculus is in danger of becoming esoteric, and those of us who teach it are in danger of being considered a *cult* if we allow eagerness for modernization to do violence to correct usage and accepted terminology.

The mathematician may treat usage with scorn, of course, as did Humpty Dumpty in his imperious reply to Alice: "When *I* use a word, it means just what

I choose it to mean—neither more nor less." But I am not convinced that mathematicians in the aggregate wish to carry this privilege to extremes. At the highest levels of research it may be that a writer can expect tolerance from his colleagues if his preoccupation is so intense that his language becomes a law unto itself. Nevertheless, a book that flouts the prevailing language of science and engineering does not constitute good text material. The student will find that functions are not exclusively in the province of mathematicians, but pervade also the business of the laboratory and the market place.

The fact of the 1960's is that mathematicians, physicists, engineers, and increasing numbers of people in the behavioral and management sciences, are accustomed to making certain types of assertions with regard to functions, such as the following:

"This function always increases."
"This function changes always at the same rate."
"This function takes only positive values."

Under these circumstances, a definition which states what a function is should be such that the essential predicate nominative may be substituted for the word "function" in each of the above sentences. In all three cases the term "variable" fits logically, whereas the term "set of ordered pairs" does not, nor does the term "correspondence," or "rule of correspondence," or "mapping." Either we must adhere to the concept that a function is a variable, or else we face the long and difficult battle of uprooting established forms of speech. In such a battle the first casualty, in my opinion, will be the reputation of calculus as a reasonable academic discipline.

As for those who doubt that the concept of a variable can be made precise, I suggest that teachers should explore this problem in a fundamental way, rather than accept a dictum. It is widely agreed, I assume, that a major purpose of calculus is to create effective models of physical situations. If it is now becoming a status-symbol to profess doubt as to what a variable is, nevertheless nature continues to go about the business of variation with carefree abandon. Antelopes run, birds fly, rockets accelerate, apples fall, nuclear bombs explode, and in general the universe seems in a state of incessant change. The invention of calculus resulted from an effort to model certain variations of nature, notably the variations of distance and velocity in a system where a body in orbit is subject to a gravitational force directed toward a center of mass.

Accordingly a characteristic task of calculus is to create a model of something that varies, and any mathematical entity that may serve this purpose (even if not so used) is properly characterized by the term "variable." Moreover, it is not difficult to define this term in a way that is compatible with the language of sets and at the same time provides a foundation for definition (A).

The language which calls $f(x)$ a function and considers it a variable proves indispensable whenever mathematicians are confronted with practical affairs, and this is noticeable at the most authoritative levels. Thus in the Report of the

Cambridge Conference on School Mathematics (Educational Services Incorporated, 1963), one finds on page 59 that a certain differential equation with a side condition "determines a unique function $f(x)$." On page 61 the same report suggests the use of Taylor's series to obtain the series for "the standard functions such as $\log(1+x)$ and $(1+x)^p$." In 1964 the William Lowell Putnam examination set the stage for a problem with the command "Let $f(x)$ be a real continuous function defined for all x," and in another problem the competitor was required to find "all continuous functions $f(x)$, for $0 \leqq x \leqq 1$, such that. . . ." If this is to be the language of professionals, why should we exclude it from the classroom?

When viewed in retrospect, the campaign to overthrow definition (A) seems an exercise in futility. The fact remains that $f(x)$, whatever one may call it, stands indispensable as the center of interest in calculus. It still has the property that educated people attribute to a variable. It is still true that, under suitable specifications, every assignment of a value to x determines uniquely a value of $f(x)$. It is still true that the English language regards a function as "any quality, trait, or fact so related to another that it is dependent upon and varies with that other." It is still true that usage, and history, and etymology, as well as the vast weight of established literature, regard $f(x)$ as a function.

Under these conditions, if the teachers and authors of calculus are supposed not to call $f(x)$ a function, then we are confronted with three (shall I say needlessly self-imposed?) problems. First, it will be necessary to find a generally accepted new name for $f(x)$; second, the established literature of science and engineering will have to be readjusted to a fundamental change in terminology; and third, we shall have to agree on what it is that shall finally be called a function. With regard to this third problem, those who would deprive $f(x)$ of its long-enduring name are not in agreement as to where to bestow the prize. Some would give the title to a set of ordered pairs, others to a correspondence, others to a rule of correspondence, others to a mapping. But none of these candidates (all admittedly important concepts) can reasonably be called a function, if the language we speak and derive from established literature is to play its proper role in mathematical education.

A STUDY OF WAYS OF HANDLING LARGE CLASSES IN FRESHMAN MATHEMATICS

V. D. TURNER, C. D. ALDERS, F. HATFIELD, HARVEY CROY, CHARLES SIGRIST, Mankato State College, Minnesota

Increased enrollments along with a desire to maintain personal contact between student and instructor make it highly desirable to explore ways of handling large classes.

The problem. Research indicates that learning in large lecture classes is not generally inferior to that in smaller lecture classes if one uses traditional achievement tests as a criterion, but experiments suggest that fewer students raise questions or interpose comments in large classes than in small. There may not

be the interaction between instructor and student necessary for a successful lecture. In view of the fact that colleges are facing doubling enrollments in the next few years, it seems imperative that research be conducted with large classes.

Experimental procedure. The study was conducted with students in Fundamentals of Mathematics (A General Education Course) and College Algebra during the winter quarter, 1963–64, at Mankato State College. A similar study was conducted with College Algebra and an integrated course of Analytic Geometry and Calculus I during the spring, 1964.

During the winter quarter, all students enrolled in Fundamentals of Mathematics and College Algebra were assigned at random to a control or to an experimental section. The control sections consisted of about fifty students each and were taught using a lecture-discussion method for each class day (four days per week for Fundamentals of Mathematics and five days per week for College Algebra). Students in the Fundamentals of Mathematics experimental section were taught by a lecture-discussion method for two days, those in College Algebra were taught by a lecture-discussion method for three days. For the remaining two class meetings, the experimental sections were randomly placed in smaller groups to receive one of three treatments. The treatment for Group I consisted of having students work together in groups of three students each with one student in the group acting as leader. The instructor in charge of the lecture was present but did a minimum of talking to the group. Group II was formed by placing the students in groups of five to six students each, with an undergraduate mathematics major in charge of each group. The undergraduate mathematics major was instructed to use a variety of methods to instruct the students. Group III was taught by a graduate assistant using a variety of methods of instruction.

During the spring quarter, the experiment was continued with College Algebra and Analytic Geometry and Calculus I. The control sections were taught as during the winter quarter (lecture-discussion for five days per week); however, the treatment for the experimental sections was modified as follows: The class received a lecture-discussion by the regular instructor for three days. For the remaining two days, the class was divided into small groups of twelve to twenty students with a senior mathematics major acting as a teaching assistant and in charge of each group. These small groups were further divided into groups with three students working together. The mathematics major was instructed to use a variety of teaching methods. The regular instructor was free to move from group to group working with individuals as he felt necessary.

For the experiment, control and experimental sections were scheduled during consecutive adjoining hours and were taught by the same instructor. The same text was used by both sections and the same tests were given each group.

Experimental results. A simple randomized design was used to test the null hypothesis of equal means for the control and experimental groups. The "F" test statistic in all classes indicated no significant differences at the 5 per cent

level. Tables showing the means of the different groups are included. The scores are those obtained on a final examination. Each row in the first table represents a different teacher who gave his own tests; hence, there is no comparison in the scores between these rows. The columns headed I, II, and III are the three groups as explained earlier. The letter n stands for the number of students in each group.

TABLES OF MEAN SCORES OBTAINED ON A FINAL EXAMINATION DURING THE WINTER QUARTER 1963–1964

Fundamentals of Math

	Control	*Group I*	*Group II*	*Group III*	*Variance Estimate*
Teacher A	$n=31$	$n=9$	$n=9$	$n=9$	357
	76	77.5	72.7	84.2	
Teacher B	$n=52$	$n=18$	$n=17$	$n=18$	341
	71	65.7	66.5	75.6	

College Algebra

Control	*Group I*	*Group II*	*Group III*	*Variance Estimate*
$n=54$	$n=13$	$n=17$	$n=17$	787
55.9	60.1	58.8	61.88	

TABLES OF MEAN SCORES OBTAINED DURING THE SPRING QUARTER 1964

College Algebra

Control	*Experimental*	*Variance Estimate*
$n=42$	$n=48$	975
$\bar{x}=126$	$\bar{x}=133$	

Calculus

Control	*Experimental*	*Variance Estimate*
$n=50$	$n=38$	205
$\bar{x}=74.3$	$\bar{x}=75.2$	

Conclusions. On the basis of the analysis and within the limitations of this study, the following conclusions were drawn:

1. The results of using the various treatments indicated no significant differences in achievement.

2. Teachers participating were able to cover the material in the experimental sections using three of five days for lecture-discussion. They were rushed in trying to cover material in two days with the four hour course.

3. The use of small groups and the use of mathematics majors as teaching assistants allowed for individual help and for active student participation.

4. The opportunity for experience provided senior mathematics majors as teaching assistants was very much worth while.

FLORIDA STATE UNIVERSITY SELECTED AS CENTER FOR RESEARCH IN COLLEGE INSTRUCTION OF SCIENCE AND MATHEMATICS

Professor Robert T. Lagemann of Vanderbilt University, Chairman of the Interim Organizing Committee (IOC) for the establishment of a Southern Center for Research in Science and Mathematics Instruction at the College Level, has announced the selection of Florida State University as the site for the regional center. The IOC has also set up a Board of Governors consisting of scientists from eighteen universities located in the southern United States, which will supersede IOC and direct the general affairs of the center.

The Interim Organizing Committee was set up at a conference attended by several dozen scientists and mathematicians which was held in New Orleans in February, 1964. This conference was sponsored by the Commission on the Undergraduate Program in Mathematics, the Commission on College Physics, the Advisory Council on College Chemistry, and the Commission on Undergraduate Education in the Biological Sciences. IOC was charged with making preliminary plans for the center, selection of a host institution, and establishment of a governing board. IOC has held a series of meetings at which these matters were considered and made the selection of the host university at its final meeting on March 8, 1966. The selection was made from a list of nine universities which had expressed an interest in serving as the host institution.

The center will engage in a variety of activities including the conduct of investigations in science and mathematics instruction; the development of new instructional materials, methods, and courses; implementation of action programs; and testing, evaluating, and disseminating its products. There will be a resident staff, and in addition, the center will make its facilities available to college and university faculty members who wish to take a leave of absence for a summer or for a year or two for conducting a specific investigation or development program.

The southern center is one of several which are being projected for establishment in various regions of the country. The centers will differ from the various Commissions on college science and mathematics, supported by the National Science Foundation, in that they will be engaged primarily in research, development, production and implementation rather than in planning and policy making.

BRIEF COMMENT

The Low Achiever in Mathematics, LAUREN G. WOODBY, Editor, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 1965 (Catalog No. FS 5.229:29061) vi+96 pgs. 35 cents.

This is a report of a conference held in Washington, D. C., March 25–27, 1964, sponsored jointly by the U. S. Office of Education and the National Council of Teachers of Mathematics. The report contains papers by a variety of people from the fields of mathematics, psychology, education and industry

on such topics as Motivation for Low Achievers, Implication of Psychological Research, Responsibilities of School Administrators, and How Can Business and Industry Cooperate With Schools on the Problem of Low Achievers in Mathematics. There is also a group of papers on current promising practices in this area.

The recommendations made by the conference are (a) to establish a National Commission of Mathematics for Low Achievers, (b) to establish research and development centers, (c) to extend the research effort—some specific directions for such research is mentioned—and (d) to develop inservice programs for teachers. Each of the above four recommendations is spelled out in some detail with specific roles and activities mentioned for the organizations whose establishment is recommended.

Guidebook to Departments in the Mathematical Sciences in the United States and Canada, The Mathematical Association of America, Buffalo, 1965.

This guidebook, prepared by the Committee on Advisement and Personnel of the Association, "is intended to provide in summary form information about the location, size, staff, library facilities, course offerings, and special features of departments in mathematical sciences in four year colleges and universities in the United States and Canada. Its purpose is to assist prospective students in these countries and abroad, in high school, junior college, or in college, and their counselors and parents in obtaining comparable information about many institutions so that the selection of a proposed place of study may be narrowed down to a few from which more detailed information may be sought on an individual basis."

The purpose of the guidebook is praiseworthy. It is unfortunate that of the 1300 questionnaires sent out only 720 were answered, since as a result the guidebook leaves unmentioned quite a few institutions. Among others, the University of Pennsylvania, Columbia University, Harvard University, and New York University are not represented. It is to be hoped that future editions of the guidebook will be more comprehensive.

Copies of the guidebook are available from the offices of the Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y., 14214, for fifty cents each.

Scientists and Engineers in Colleges and Universities, 1961, Surveys of Science Resources Series, NSF, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. (Catalog No. NSF 65-8) xi+125 pgs. 65 cents.

"This report summarizes the final results of the National Science Foundation's survey of scientific and engineering personnel employed in institutions of higher education in the United States in 1961. The Foundation has published preliminary findings of this survey as well as final reports on two previous surveys covering employment of such personnel for the academic years 1953–54 and 1957–58."

"The employment of scientists and engineers in 1712 institutions of higher education covered in this report totaled 175,600 in March 1961. Of these, 108,100 were faculty members, 35,000 were employed graduate students, and the remaining 32,500 were listed as other professional personnel. The field-of-science distribution of scientists and engineers was as follows: 36 percent in the life sciences, 28 percent in the physical sciences, 16 percent in engineering, 15 percent in the social sciences, and 5 percent in psychology." Of the 175,600 individuals covered by the report 14,800 were mathematicians; of these 10,490 were faculty members, 2780 employed graduate students, and 1530 were other professional personnel.

Mathematics in Colleges and Universities—A Comprehensive Survey of Graduate and Undergraduate Programs (Final Report), CLARENCE B. LINDQUIST, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C. 1965 (Catalog No. FS 5.256:56018) xi+104 pgs. 60 cents.

"The survey of mathematics programs reported here was the most comprehensive depth study of programs within a discipline ever undertaken in the United States. A total of 877 colleges and universities granting bachelor's or higher degrees, or about 82 percent of the 1069 to which questionnaires were sent, responded. Both undergraduate and graduate programs were surveyed, and information was solicited on curriculums, degrees, course offerings, enrollments, credit requirements, examination requirements, special features, innovations, and trends."

The report contains narrative material and numerous charts giving statistical data. To give a notion of the sorts of material included there follows a partial list of section headings: Entering Freshman Levels in Mathematics, Characteristics of Curriculums for the Mathematics "Major," Course Offerings and Enrollments, Prerequisite Instruction, Required Admissions Examinations that Included Mathematics, Placement Examinations in Mathematics, Programs of Advanced Standing in Mathematics, Undergraduate Honors Programs in Mathematics (including a detailed account of programs at Carleton, Wesleyan, Dartmouth, and the University of Maryland), and so on. There is also some study of activities designed to stimulate interest in mathematics such as mathematics clubs, nonstandard teaching techniques, inservice education of mathematics teachers, and library organization and computers at colleges and universities.

Mathematics Curriculum: New Study, E. G. BEGLE, Letter to the Editor of *Science*, February 11, No. 3711, 151(1966) 632.

Despite the innovations in school mathematics introduced in the last eight years since the founding of SMSG in 1958 it is suggested in this letter that "longer-range planning and experimentation should be started before present materials become frozen into a newly orthodox pattern that will require another upheaval a few years hence." The board of SMSG "has, therefore, decided to

convene a group to design a new sequential curriculum for grades 7 to 12 and to plan appropriate experimental materials. Major emphasis is to be given to the design of courses which exploit recent progress and to a sequential curriculum which will be responsive to the rapidly developing needs for mathematics in our society." Panels will be convened to carry out the above decision, and a list of questions which the panels will consider is given.

An article in the Sunday New York Times for February 20, 1966, on page 92, "New Math Fears it is Growing Old," by Harry Schwartz, reviews the above letter.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J., 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to M. S. Klamkin, Ford Scientific Laboratory, P.O. Box 2053, Dearborn, Mich. 48121. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before January 31, 1967.

Editorial Announcement. **We regret that we cannot be responsible for any solution to any Elementary Problem, submitted on or after October 1, 1966, unless it was directed to M. S. Klamkin at the foregoing address.**

E 1905. *Proposed by R. L. Graham and L. A. Shepp, Bell Telephone Laboratories.*

Let $x_1=x$, $x_{n+1}=x^{x_n}$ for $n=1, 2, \cdots$. If $a>e^{1/e}$, prove there is a γ for which $\lim_{n\to\infty} \gamma_n/a_{n+1}$ exists and is a positive number. Is γ unique?

E 1906. *Proposed by Erwin Just, Bronx Community College*

Prove that

$$\sum_{k=1}^{2n} \tau(k) - \sum_{k=1}^{n} [2n/k] = n,$$

where $\tau(n)$ is the number of divisors of n and $[x]$ is the greatest integer not exceeding x.

E 1907. *Proposed by Sidney Spital, California State Polytechnic College*

Steffenson's method for solving the equation $x=f(x)$ consists of forming recursively a sequence $\{x_n\}$, using

$$x_{n+1} = x_n - \frac{[f(x_n) - x_n]^2}{f(f(x_n)) - 2f(x_n) + x_n},$$

which (hopefully) converges to a solution s. Show that if f is twice continuously differentiable and if $f'(s)\neq 1$, then the errors $e_n = x_n - s$ satisfy the following condition of quadratic convergence:

$$e_{n+1} = \frac{f''(s)f'(s)}{2(f'(s) - 1)} e_n^2 + O(e_n^3).$$

E 1908. *Proposed by Edgar Karst, University of Oklahoma*

Four equal spheres are inscribed in a hemisphere of horizontal base such that each sphere touches two others and is tangent to both the hemisphere and its base. A radial section of the figure is now made by a vertical plane passing through the center of the hemisphere and the center of one of the inscribed spheres, and in this section a radius is drawn tangent to the circular section of the inscribed sphere. We now have a circular sector with its inscribed circle. Prove that the little circles inscribed in the corners of the sector and externally tangent to the inscribed circle are equal to one another.

E 1909. *Proposed by C. C. Lindner, Coker College, Hartsville, S. C.*

Prove that the center of a group is properly contained in every maximal subgroup having composite index.

E 1910. *Proposed by R. L. Graham, Bell Telephone Laboratories*

Show that there exists a sequence $a_1 < a_2 < \cdots$ of integers such that every positive integer occurs uniquely as a difference $a_i - a_j$ for some i and j. What can be said about the growth of such a sequence?

E 1911. *Proposed by B. R. Toskey, Seattle University*

Suppose $a_{11}, a_{12}, \cdots, a_{1n}$ are given integers whose greatest common divisor is 1, and suppose $n\geqq 2$. Is it always possible to find a matrix (a_{ij}) with the given integers in the first row and all a_{ij} integers such that $\det(a_{ij})=1$?

E 1912. *Proposed by R. E. Williamson, Dartmouth College*

Prove the following generalization of the Pythagorean relation. Let P be a parallelotope spanned by k vectors in R^n, $k\leqq n$, that is, the set of combinations $a_1V_1 + \cdots + a_kV_k$, where $V_1, \cdots, V_k$ are the given vectors and the a_i satisfy

$0 \leqq a_i \leqq 1$, $i=1, \cdots, k$. Let $\{P_j\}$, $j=1, 2, \cdots, \binom{n}{k}$, be the projections of P into the k-dimensional coordinate flats of R^n, $\binom{n}{k}$ in number. Then

$$V_k^2(P) = \sum_{j=1}^{\binom{n}{k}} V_k^2(P_j),$$

where V_k is k-dimensional volume in R^n.

E 1913. *Proposed by R. A. Jacobson, South Dakota State University*

Given the set of $2n$ integers $\{\pm a_1, \pm a_2, \cdots, \pm a_n\}$ and a positive integer $m < 2^n$. Show that it is always possible to select a subset S such that (i) $\pm a_i$ are not both contained in S; (ii) the sum of the elements in S is divisible by m.

E 1914. *Proposed by Don Redmond, Fremont, California*

Prove that the nth roots of unity constitute the only set of nth roots of a complex number which is a group under multiplication.

SOLUTIONS OF ELEMENTARY PROBLEMS

$YY' = XCC'X'$

E 1787 [1965, 544]. *Proposed by Kenneth Kloss, National Bureau of Standards*

Prove the identity

$$\sum_{1 \leq i \leq j \leq n} x_i x_j = \sum_{k=1}^{n} \frac{k+1}{2k}\left(x_k + \frac{x_{k+1} + \cdots + x_n}{k+1}\right)^2.$$

I. *Solution by D. A. Zave, Case Institute of Technology (student).* For $i \leqq j$, it is easily seen that the coefficient of $x_i x_j$ in the kth term of the right member is

$$\begin{array}{ccc} 0 & \text{if} & i < k, \\ (k+1)/2k & \text{if} & i = j = k, \\ 1/2k(k+1) & \text{if} & k < i = j, \\ 1/k & \text{if} & k = i < j, \\ 1/k(k+1) & \text{if} & k < i < j. \end{array}$$

Hence, the coefficient of x_i^2 in the sum on the right is

$$\sum_{k=1}^{i-1} \frac{1}{2k(k+1)} + \frac{i+1}{2i} = \frac{1}{2}\sum_{k=1}^{i-1}\left(\frac{1}{k} - \frac{1}{k+1}\right) + \frac{i+1}{2i} = 1,$$

while the coefficient of $x_i x_j$ $(i<j)$ is $\sum_{k=1}^{i-1} 1/(k(k+1)) + 1/i = 1$.

II. *Solution by D. C. B. Marsh, Colorado School of Mines.* Treating the problem as a linear transformation and employing matrix notation, we let

$X=(x_1, x_2, \cdots, x_n)$, $Y=(y_1, y_2, \cdots, y_n)$, and

$$y_k = \left(\frac{k+1}{2k}\right)^{1/2}\left(x_k + \frac{x_{k+1} + \cdots + x_n}{k+1}\right)$$

for $k=1, 2, \cdots, n$. The right-hand expression may thus be written as $YY' = (XC)(C'X')$, where $C=(c_{ij})$ is given by

$$c_{ij} = \begin{bmatrix} 0 & \text{if} \quad i < j, \\ \left(\dfrac{j+1}{2j}\right)^{1/2} & \text{if} \quad i = j, \\ \left(\dfrac{j+1}{2j}\right)^{1/2}\dfrac{1}{j+1} & \text{if} \quad i > j. \end{bmatrix}$$

By routine matrix multiplication we compute $CC'=P=(p_{ij})$: For $i>j$ (and, by symmetry of CC', for $i<j$),

$$p_{ij} = \sum_{t=1}^{j} c_{it}c_{jt} = \sum_{t=1}^{j-1} \frac{1}{2t(t+1)} + \frac{1}{2j} = \frac{1}{2};$$

$$p_{ij} = \sum_{t=1}^{j-1} 1/2t(t+1) + (j+1)/2j = 1, \qquad \text{for } i = j.$$

Thus, the right-hand expression is equivalent to XPX'—the desired left-hand quantity.

III. *Solution by R. E. Maas, University of Santa Clara.* Denoting the right-hand side of the purported equality by R_n, we have for $n>1$,

$$\begin{aligned} R_n - R_{n-1} &= \frac{n+1}{2n}x_n^2 + \sum_{k=1}^{n-1}\frac{k+1}{2k}\left[\left(x_k + \frac{x_{k+1}+\cdots+x_n}{k+1}\right)^2 \right. \\ &\qquad \left. - \left(x_k + \frac{x_{k+1}+\cdots+x_{n-1}}{k+1}\right)^2\right] \\ &= \frac{n+1}{2n}x_n^2 + \sum_{k=1}^{n-1}\frac{k+1}{2k}\left[\left(\frac{x_n}{k+1}\right)^2 \right. \\ &\qquad \left. + \frac{2x_n}{k+1}\left(x_k + \frac{x_{k+1}+\cdots+x_{n-1}}{k+1}\right)\right] \\ &= \frac{n+1}{2n}x_n^2 + \frac{x_n^2}{2}\sum_{k=1}^{n-1}\frac{1}{k(k+1)} + x_n\sum_{k=1}^{n-1}\frac{x_k}{k} \\ &\qquad + x_n\sum_{k=1}^{n-2}\sum_{j=k+1}^{n-1}\frac{x_j}{k(k+1)} \end{aligned}$$

$$= x_n^2 + x_n \sum_{k=1}^{n-1} \frac{x_k}{k} + x_n \sum_{j=2}^{n-1} x_j \sum_{k=1}^{j-1} \frac{1}{k(k+1)}$$

$$= x_n \sum_{k=1}^{n} x_k = \sum_{1 \leq i \leq j \leq n} x_i x_j - \sum_{1 \leq i \leq j \leq n-1} x_i x_j,$$

where we have used the equality $\sum_{k=1}^{r-1} 1/k(k+1) = 1 - 1/r$. The desired result now follows by induction.

Also solved by A. N. Aheart, Tim Betts, E. O. Buchman, Orin Chein, W. O. Egerland, Simon Green, S. H. Greene, D. M. Hancasky, Stephen Hoffman, Edward Hook, E. S. Langford, Agnis Kaugars, M. S. Klamkin, Richard Loebl, Gus Mavrigian (2 solutions), Y. Mayer ben-David, Robert Patenaude, C. B. A. Peck, B. V. Rao (India) (2 solutions), Simeon Reich (Israel), B. E. Rhoades, P. A. Scheinok, Klaus Schmitt, P. S. Schnare, Michael Sheridan, Al Somayajulu, Sidney Spital, Guy Torchinelli, Elias Toubassi, W. C. Waterhouse, D. R. Wilder, David Zeitlin, and the proposer.

A Curious Subset of (0, 1)

E 1788 [1965, 544]. *Proposed by Joel Pitcairn, Bryn Athyn, Pa.*

Let J be the open interval $\{t: 0<t<1\}$. We call a subset of J a C-set in case it contains $1-t$ and ts whenever it contains t and s.

(1) Let E be a C-set. Prove (a) if E is nonvoid, then E is dense in J; and (b) if $\text{int}(E)$ is nonvoid, then $E=J$.

(2) Write t' for $1-t$. Suppose X is a real vector space, $A \subset X$, and f a real-valued function on A. (a) Let E be the set of all those $t \in J$ such that $x, y \in A$ implies $tx+t'y \in A$. (Thus A is midpoint convex iff $1/2 \in E$.) Prove that E is a C-set. (b) Suppose A is convex, and let E be the set of those $t \in J$ such that $x, y \in A$ implies $f(tx+t'y) \leqq tf(x)+t'f(y)$. Prove that E is a C-set.

Solution by G. P. Speck, U. S. Naval Academy, Annapolis, Md. (1a) Suppose $t \in E$. Then $(1-t^n)^m \in E$ for $n=1, 2, 3, \cdots$ and $m=1, 2, 3, \cdots$. This collection, $\{(1-t^n)^m\}$, is easily seen to be dense in J.

(1b) If $(s, t) \subset E$, then $(s^n, t^n) \subset E$ and hence $(1-t^n, 1-s^n) \subset E$. From this we see that for any $r \in J$ there exists $(a, b) \subset E$ such that $r<a$. Now from (1a), there exists $x \in E$ such that $r/b<x<r/a$; i.e., $xa<r$ and $xb>r$. Therefore, there exists $c \in (a, b)$ such that $xc=r$, and hence $r \in E$.

(2a) Consider $x, y \in A$. Then $ty+(1-t)x \in A$ if $t \in E$, which implies $1-t \in E$. Also, $tx+(1-t)y, y \in A$ if $t \in E$. Thus, if $s \in E$, then $s(tx+(1-t)y)+(1-s)y \in A$; i.e., $stx+(1-st)y \in A$, which implies $st \in E$.

(2b) We are to show that $s, t \in E$ implies $1-t$, $st \in E$. Consider $x, y \in A$. Then $f((1-t)x+ty)=f(ty+(1-t)x) \leqq tf(y)+(1-t)f(x)$, which implies $1-t \in E$. Also, since $y \in A$ and, by convexity, $tx+(1-t)y \in A$, then:

$$\begin{aligned} f[s(tx+(1-t)y)+(1-s)y] &\leqq sf(tx+(1-t)y)+(1-s)f(y) \\ &\leqq s[tf(x)+(1-t)f(y)]+(1-s)f(y) \\ &= stf(x)+(1-st)f(y), \end{aligned}$$

which implies $st \in E$.

Also solved by E. O. Buchman, N. J. Fine, E. S. Langford, and the proposer. Partial solution by D. M. Hancasky,

A Special Case of a Theorem of H. Whitney

E 1789 [1965, 544]. *Proposed by R. A. Bell, Kansas City, Mo.*

Suppose that $g(x)$ has its first $n+1$ derivatives defined and continuous in $[-1, 1]$. Define $y(x)=g(x)/x$ for $x\neq 0$ and $y(0)=g'(0)$. If $g(0)=0$, prove that $y^{(n)}(0)=d^ny/dx^n|_{x=0}$ exists and equals $g^{(n+1)}(0)/(n+1)$.

I. *Solution by W. O. Egerland, USA Nuclear Defense Laboratory, Edgewood Arsenal.* For $x\neq 0$ we have

$$xy(x) = g(x) = \int_0^x g'(t)dt \quad \text{or} \quad y(x) = \int_0^1 g'(xu)du.$$

Since $g'(xu)$ is $(n-1)$ times continuously differentiable with respect to x in the rectangles $R_1[0\leqq u\leqq 1,\ 0<x_1\leqq x\leqq 1]$ and $R_2[0\leqq u\leqq 1,\ -1\leqq x\leqq x_2<0]$, we obtain by Leibniz' Rule

$$y^{(n-1)}(x) = \int_0^1 u^{n-1}g^{(n)}(xu)du \qquad (x\neq 0)$$

so that

$$\frac{1}{x}\left(y^{(n-1)}(x) - g^{(n)}(0)/n\right) = \frac{1}{x^{n+1}}\int_0^x t^{n-1}[g^{(n)}(t) - g^{(n)}(0)]dt \to \frac{g^{(n+1)}(0)}{n+1} = y^{(n)}(0)$$

as $x\to 0+$ or $x\to 0-$, by L'Hospital's Rule.

The proof shows that only the existence of $g^{(n+1)}(x)$ at $x=0$ is needed to establish the assertion of the problem.

II. *Solution by M. S. Klamkin, Mathematical and Theoretical Sciences Scientific Laboratory, Ford Motor Company, Dearborn, Michigan.* Let

$$D^m[g(x)/x] = F_m(x)/x^{m+1},$$

so that $DF_0(x)=D\{xD^0[g(x)/x]\}=Dg(x)=x^0Dg(x)$. Assume that

$$(*) \qquad DF_m(x) = x^mD^{m+1}g(x).$$

Then

$$\frac{F_{m+1}(x)}{x^{m+2}} = D^{m+1}\frac{g(x)}{x} = D\frac{F_m(x)}{x^{m+1}} = \frac{xDF_m(x) - (m+1)F_m(x)}{x^{m+2}},$$

whence $F_{m+1}(x)=xDF_m(x)-(m+1)F_m(x)$ and

$$\begin{aligned} DF_{m+1}(x) &= xD^2F_m(x) - (m+1)DF_m(x) \\ &= xD[x^mD^{m+1}g(x)] - (m+1)x^mD^{m+1}g(x) \\ &= x^{m+1}D^{m+2}g(x). \end{aligned}$$

It follows by induction that (*) is true for all nonnegative integers m.

By L'Hospital's Rule, $\lim_{x\to 0} D^n[g(x)/x]$ will exist if $\lim_{x\to 0} [DF_n(x)/Dx^{n+1}]$ exists. By (*), this latter limit is $g^{(n)}(0)/(n+1)$.

III. *Comment by L. E. Pursell, Grinnell College.* This problem is a special case of a theorem proved by Hassler Whitney, *Differentiability of the remainder term in Taylor's formula*, Duke Math. J., 10 (1943) 153–158, concerning the derivatives of the "remainder quotients," $R_{n-1}(x)/x^n$, where $R_{n-1}(x)=g(x)-\sum_{k=0}^{n-1} g^{(k)}(0)x^k/k!$. Professor Whitney also extends his results to functions of several variables. Whitney's theorem for functions of one variable also appears as a problem in Watson Fulks, *Advanced Calculus: An Introduction to Analysis*, Wiley, 1961, p. 140, problem C1.

I devised a proof similar to that in Solution I above several years ago after encountering the integral transform $\int_0^1 g'(xt)dt$ in K. Nomizu, *Lie groups and differential geometry*, Math. Soc. of Japan, 1956, p. 7. A similar proof for a slightly different theorem appears in Sigurdur Helgason, *Differential Geometry and Symmetric Spaces*, Academic Press, 1962, pp. 9–10.

Also solved by E. O. Buchman, J. L. Gieser, Richard Gisselquist, D. M. Goldschmidt, R. W. Hansell, Stephen Hoffman, S. C. King, E. S. Langford, J. C. Lazzara, D. C. B. Marsh, Y. Mayer ben-David, Norman Miller, J. M. Perry, Simeon Reich (Israel), Al Somayajulu, Sidney Spital, H. H. Wong, and the proposer. Partial solutions by P. J. Campbell, G. A. Fisher, D. M. Hancasky, J. C. Hickman, D. R. Lehman, N. T. Sheth, R. Sivaramakrishnan (India), K. L. Yocom, and David Zeitlin.

Some of those listed as partial solvers obtained $y^{(k)}(x)$ by differentiating Taylor's Formula throughout but forgot that the c in the remainder $x^{n+1}g^{(n+1)}(c)/(n+1)!$ is also a function of x. Others assumed that $xy^{(n+1)}(x)$ has limit 0 as $x\to 0$.

A Family of Catenaries and Its Envelope

E 1790 [1965, 544]. *Proposed by Alvin Hausner, City College of New York*

For what real numbers m, $n\neq 0$ do there exist functions $y=f(x)$, positive and having continuous derivatives for all real x, such that the mth power of the length of arc of the curve $y=f(x)$ between any two points is proportional to the nth power of the area under the curve between these points? Determine the functions in those cases where they exist.

Solution by Louis Padulo, Student, Georgia Institute of Technology. This is a special case of the

PROBLEM. If $F(x, y_1, y_2, \cdots, y_k)$ and $G(x, y_1, y_2, \cdots, y_k)$ are continuous and positive for $a<x<b$ and $-\infty<y_i<\infty$ $(i=1, 2, \cdots, k)$, for what real numbers m and n and functions $y=y(x)$ with continuous kth derivatives is

$$(*)\quad \left[\int_\alpha^x F(t, y, y', \cdots, y^{(k)})\,dt\right]^m = A\left[\int_\alpha^x G(t, y, y', \cdots, y^{(k)})dt\right]^n$$

for $a<\alpha<x<b$ and A a positive constant?

SOLUTION. If $m<n$, divide by $(x-\alpha)^m$ and let $x\to\alpha+$. By the Fundamental Theorem of the Integral Calculus, the left member has limit

$$F(t, y, y', \cdots, y^{(k)})\Big|_{t=\alpha}$$

and the right member, zero . . . contrary to F's being positive. Thus, $m\geqq n$ and, similarly, $m\leqq n$. Consequently, we must have $m=n$ and $\int_\alpha^x F dt=B\int_\alpha^x G dt$ with $B=\sqrt[n]{A}$. Differentiating with respect to x, y must satisfy the differential equation $F(x, y, y', \cdots, y^{(k)})=BG(x, y, y', \cdots, y^{(k)})$.

For the problem posed, $k=1$, $f(x, y, y')=\sqrt{(1+y'^2)}$, and $G(x, y, y')=y$, and the corresponding differential equation is known to have the solutions

$$y = B^{-1}\cosh B(x-c) \quad \text{and} \quad y = 1/B.$$

That these are indeed solutions of (*) is easily verified.

Also solved by E. O. Buchman, P. J. Campbell, Michael Goldberg, M. S. Klamkin, E. S. Langford, D. R. Lehman, D. C. B. Marsh, Sidney Spital, and the proposer.

The Cardioid as a Union of Circles

E 1791 [1965, 544]. *Proposed by R. G. Buschman, State University of New York at Buffalo*

(a) Consider the set S of circles such that (i) the center of every circle of S is interior to the unit circle and (ii) the point (1, 0) is not interior to any circle of S. Show that the union of the interiors of the circles of the set S is the interior of a cardioid.

(b) Consider the set S' such that (i) holds and (ii) is replaced by (ii′) none of the points representing the nth roots of unity are interior to any circle of S'. What figure replaces the cardioid of (a)?

Solution by D. C. B. Marsh, Colorado School of Mines.

(a) We let (a, b) vary over the interior of the unit circle and find all (x, y) satisfying $(x-a)^2+(y-b)^2\leqq(a-1)^2+b^2$. The desired region is bounded by the envelope of the maximal circles of this set. This envelope is routinely computed to be $(x^2+y^2-1)=2\{(x-1)^2+y^2\}^{1/2}$ which becomes, for $x-1=r\cos\theta$ and $y=r\sin\theta$, $r=2(1-\cos\theta)$—a cardioid.

(b) If we label the origin as O, the point (1, 0) as U, and the points $(\cos\theta, \sin\theta)$ as P_θ, the circle with center at P_θ and radius $P_\theta U$ cuts the line through U and parallel to OP_θ in a point $P_\theta^*(\neq U)$ which is on the aforementioned cardioid. Circles drawn with centers (a, b) and radii extending to U will not include any of the other nth roots of unity iff (a, b) is within the circular sector bounded by arc $P_{-\pi/n}P_{\pi/n}$ and radii $OP_{-\pi/n}$ and $OP_{\pi/n}$. Let us label as C the corresponding section of the cardioid: $P^*_{-\pi/n}UP^*_{\pi/n}$. The region sought in (b) has boundary consisting in part of arcs congruent to C, symmetric about the lines $OP_{2k\pi/n}$ $(k=0, 1, \cdots, n-1)$, concave towards O, and with cusps at $P_{2k\pi/n}$; the endpoints of consecutive arcs are necessarily joined by circular arcs with centers at $P_{(2k-1)\pi/n}$, radii of $2\sin(\pi/2n)$ and concave towards O.

Also solved by E. O. Buchman, Michael Goldberg, M. S. Klamkin, Robert Patenaude, J. M. Quoniam (France), and the proposer.

Continuity from the Darboux Property and the Closedness of $g^{-1}(r)$

E 1792 [1965, 544]. *Proposed by A. M. Gleason, Harvard University*

Let f be a function from the reals to the reals, differentiable at every point. Suppose that, for every r, the set of points x, where $f'(x)=r$, is closed. Prove that f' is continuous.

Solution by N. J. Fine, Pennsylvania State University. It is well known that an everywhere existing derivative possesses the Darboux (Intermediate Value) Property:

1. *For every r, if $g(x)<r<g(y)$, there is a z between x and y such that $g(z)=r$.*

The stated result is therefore a corollary of the following

THEOREM. *Let g be a function from the reals to the reals satisfying* (1) *and*
2. *For every r, $\{x: g(x)=r\}$ is closed. Then g is continuous.*

Proof. Let $y_n \to x_0$, with $g(x_0)<r<g(y_n)$. By (1), there is a sequence $z_n \to x_0$ with $g(z_n)=r$. Thus x_0 is in the closure of $\{x: g(x)=r\}$. By (2), $g(x_0)=r$, a contradiction. Hence $\overline{\lim}\, g(y_n) \leqq g(x_0)$. Similarly, $\underline{\lim}\, g(y_n) \geqq g(x_0)$, so $\lim g(y_n) =g(x_0)$, and g is continuous.

Note that (1) and (2) need only be assumed for the same dense set of r's.

Also solved by E. O. Buchman, B. J. Cerimele, O. W. Dixon, W. G. Dotson, Jr., Vlaicu Drilea (Nigeria), W. B. Galvin, D. M. Goldschmidt, Harry Guess, R. W. Hansell, Horst-E. Lahmann (Germany), Joe Lipman, Y. Mayer ben-David, W. J. Pervin, Al Somayajulu, G. P. Speck, Benjamin Volk, W. C. Waterhouse, K. L. Yocom, D. S. Zave, and the proposer.

Pervin calls attention to C. H. Rowe, Bull. Amer. Math. Soc. 32, (1926) 285–287. Drilea points out that the result here is a special case of problem 18, page 131, in H. L. Royden, *Real Analysis.*

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before March 31, 1967.

5397 [1966, 548]. *Correction.* Replace $I_\nu(t)$ by $I_{-\nu}(t)$.

5410. *Proposed by H. W. Guggenheimer, University of Minnesota*

Prove that $x^{k+1}y(x^{-k})=y(x)$, $x>0$, has no analytic nontrivial ($y(x)\neq x$) solution for $k\neq 1$. (This functional equation was proposed by John Bernoulli in 1696, but no valid discussion seems to have been given.)

5411. *Proposed by Azriel Rosenfeld, University of Maryland*

Let G be a group every subgroup of which is a direct factor of G. Prove that G is a direct product of cyclic groups of prime orders.

5412. *Proposed by Paul Erdös, University of Illinois*

Let $a_1 < a_2 < \cdots$ be an infinite sequence of integers. Put A_n for the least common multiple of $a_1, a_2, \cdots, a_n$. Prove that $A_n > n^{c_1 \log \log n}$, but that there is a sequence $a_1 < a_2 < \cdots$ so that $A_r < n^{c_2 \log \log n}$.

5413. *Proposed by J. L. Selfridge, Pennsylvania State University and the University of Illinois, and Paul Erdös.*

Let $a_1 < a_2 < \cdots$ be an infinite sequence of integers. Denote by $A_n^{(k)}$ the least common multiple of $a_n, a_{n+1}, \cdots, a_{n+k-1}$. Prove that the number of indices n for which $A_n^{(k)} < x$ is less than $cx^{1/k}$.

5414. *Proposed by L. Carlitz, Duke University*

Show that

$$(1) \quad \int_0^1 \int_0^1 \{(1-x^2)(1-y^2)(1-x^2y^2)\}^{-1/2} dxdy = \frac{(\Gamma(\frac{1}{4}))^4}{16\pi}.$$

$$(2) \quad \int_0^1 \int_0^1 \{(1-x^2)(1-y^2)(1+x^2y^2)\}^{-1/2} dxdy = \frac{\pi^2}{4}\left\{\frac{\Gamma(9/8)}{\Gamma(5/4)\Gamma(7/8)}\right\}^2.$$

5415. *Proposed by R. D. Driver, Sandia Corporation*

The ordinary differential equation

$$x'(t) = 1 - g(x(t)) + g(t) \qquad \text{for } t > 0 \text{ with } x(0) = 0$$

has a solution $x(t) = t$. Is the solution unique (locally) if $g(t)$ is continuous for $t \geqq 0$?

(The proposer has a partial solution: There is no such $g(t)$ for which there exists a solution $x(t) < t$ for $t > 0$ with $x'(t)$ nonincreasing. The complete solution is desired.)

5416. *Proposed by W. J. Leahey, University of Illinois*

Let K be a number field and θ a primitive element for K/Q, i.e. $K = Q(\theta)$, where Q is the field of rational numbers. Let $f(X)$ be the minimal polynomial of θ over Q and denote by $f'(X)$ the derivative of $f(X)$. Is $f'(\theta)$ necessarily a primitive element for K/Q?

5417. *Proposed by M. V. Chari and John Slivka, State University of New York at Buffalo*

Show that $\int \cdots \int_S (y-z)dx_1 \cdots dx_r = m-n$, where $y = \min_i\{m, x_i + \frac{1}{2}\}$, $z = \max_i\{n, x_i - \frac{1}{2}\}$, $n < m$, and

$$S = \{(x_1, \cdots, x_r): n - \tfrac{1}{2} < x_i < m + \tfrac{1}{2} \text{ for all } i (i = 1, \cdots, r) \text{ and } \max_i\{x_i\} - \min_i\{x_i\} < 1\}.$$

5418. *Proposed by A. A. Mullin, University of California, Livermore*

Let ζ be Riemann's zeta-function and p a positive prime number. Determine the value $\Pi_p\zeta(p)$. (Hint: $2.05<\Pi_p\zeta(p)<2.08$.) Let c be a positive composite number and determine $\Pi_c\zeta(c)$.

5419. *Proposed by F. D. Faulkner, U. S. Naval Postgraduate School*

Let A_{nn} be a real, symmetric, positive semi-definite matrix. Let A_{pn} be the matrix obtained from A_{nn} by omitting the last $n-p$ rows, and A_{pp} be obtained by omitting the last $n-p$ columns of A_{pn}. Prove that rank A_{pp} = rank A_{pn}.

SOLUTIONS OF ADVANCED PROBLEMS

Zones of Equal Thickness and Equal Area

5305 [1965, 674]. *Proposed by Frank Dapkus, Seton Hall University, South Orange, N. J.*

Is it true that of all surfaces of revolution only spheres and circular cylinders have the property that areas of zones of equal thickness are equal?

Solution by Michael Goldberg, Washington, D. C. In the statement of the problem, the inclusion of the cylinder implies that the zones are to be taken as normal to the axis of the surface of revolution.

If the area of a zone is dA, and its thickness is dy, then dA/dy is constant. If the radius is x, then

$$dA = 2\pi x(dx^2 + dy^2)^{1/2},\ dA/dy = 2\pi x((dx/dy)^2 + 1)^{1/2} = k,$$

$$(dx/dy)^2 = (k^2 - 4\pi^2x^2)/4\pi^2x^2.$$

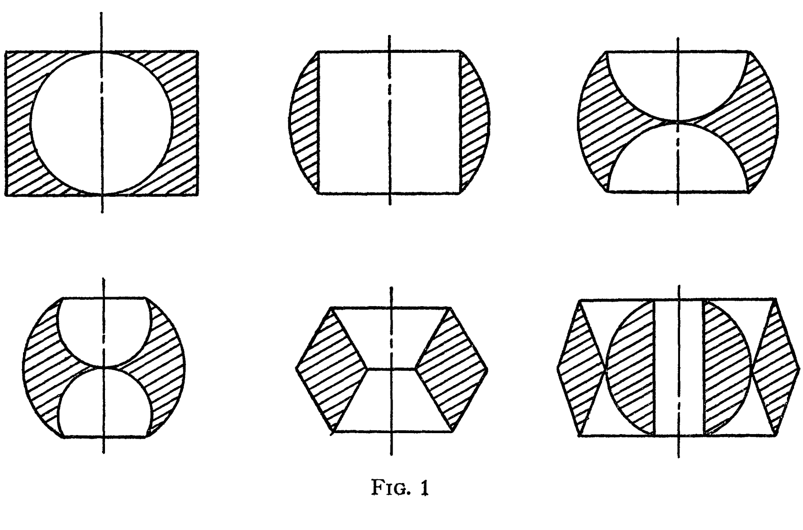


FIG. 1

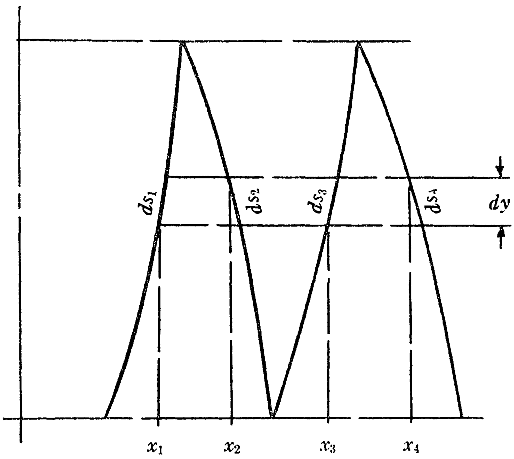


FIG. 2

If $k^2-4\pi^2x^2\neq 0$, then

$$dy = 2\pi x dx/(k^2 - 4\pi^2x^2)^{1/2} = x dx/((k/2\pi)^2 - x^2)^{1/2}.$$

Hence $y=(R^2-x^2)^{1/2}+c$, where $R=k/2\pi$, or $x^2+(y-c)^2=R^2$. This is a circle of radius R, and it generates a sphere.

However, if $k^2-4\pi^2x^2=0$, then $x^2=R^2$, and this generates a cylinder.

If permissible surfaces of revolution include nonconvex surfaces and surfaces of higher genus, then we may have combinations of co-axial cylinders, spheres, and pairs of similar cones as shown in Fig. 1.

A more general class of solutions is obtained by the use of a set of $n-1$ arbitrary curves, to which is added an nth curve satisfying the equation

$$2\pi(x_1ds_1 + x_2ds_2 + \cdots + x_{n-1}ds_{n-1} + x_nds_n) = kdy,$$

as shown in Fig. 2.

Also solved by P. N. Bajaj, J. S. Haines & G. G. Goethals, D. A. Hejhal, Norman Miller, C. E. Olson, E. J. F. Primrose (England), Fred & Judith Richman, Simon Vatriquant (Belgium), and the proposer.

Sidney Spital points out that the solution in the instance of a surface of revolution generated by a continuously differentiable function may be found in Mathematics Magazine, v. 38, no. 3, p. 183.

Infinitely Many Roots for sin $\phi(x)=\psi(x)$

5306 [1965, 674]. *Proposed by Fred Gross, National Bureau of Standards*

Let $\phi(x)$ and $\psi(x)$ be any two entire functions with $\phi(x)$ nonconstant. If $\psi(x)$ is of finite order and in particular of order less than k whenever $\phi(x)$ is a polynomial of degree k, then $\sin\phi(x)=\psi(x)$ must have infinitely many solutions. (Compare 5102 [1964, 564].)

Solution by the proposer. We use the following theorem of Borel (see e.g. E. Borel, *Leçons sur les Fonctions Entières*, 1921, Ch. V): Let $a_i(z)$ be an entire function of order at most ρ, let $g_i(z)$ also be entire and let $g_i(z)-g_j(z)$ $(i\neq j)$ be a transcendental function or polynomial of degree higher than ρ. Then

$$\sum_{i=1}^{n} a_i(z)e^{g_i(z)} = a_0(z)$$

holds only when the following conditions are satisfied

$$a_0(z) = a_1(z) = \cdots = a_n(z) = 0.$$

Assume that $\sin \phi(z)-\psi(z)$ has at most a finite number of solutions. We may then write

$$\sin \phi(z) - \psi(z) = \frac{e^{i\phi(z)}}{2i} - \frac{e^{-i\phi(z)}}{2i} - \psi(z) = Q(z)e^{\gamma(z)},$$

where $Q(z)$ is a polynomial and $\gamma(z)$ is entire. If neither $i\phi(z)+\gamma(z)$ nor $i\phi(z)-\gamma(z)$ is a polynomial then our conclusion follows from Borel's theorem. Otherwise, assume that $i\phi(z)+\gamma(z)$ is a polynomial; we need only consider the case when it is of degree less than or equal to the order of $\psi(z)$. Dividing by $e^{\gamma(z)}$ we get

$$\frac{e^{i\phi(z)-\gamma(z)}}{2i} - \frac{e^{-i\phi(z)-\gamma(z)}}{2i} - \psi(z)e^{-\gamma(z)} = Q(z).$$

Since $\phi(z)$ cannot be a polynomial of degree less than or equal to the order of $\psi(z)$, Borel's theorem again applies, and it follows that $i\phi(z)-\gamma(z)=-\gamma(z)+c$, where c is a constant. This renders $\phi(z)$ a constant contrary to the hypothesis.

Real Functions Arbitrarily Defined on a Countable Set

5308 [1965, 674]. *Proposed by Louis Comtet, Boulogne, France*

Let $A=\{a_1, a_2, \cdots\}$, $B=\{b_1, b_2, \cdots\}$ be two sequences of real numbers with the a_i all different and the $b_i\neq 0$. Define the function $f(x)$ by $f(x)=b_n$ if $x=a_n$ and $f(x)=0$ if $x\notin A$. (1) Show that $f(x)$ is of class $\leqq 2$ (of Baire). (2) Is it true that the two following conditions are equivalent?

(i) $\lim_{n\to\infty} b_n=0$,
(ii) $f(x)$ belongs to class 1.

Solution by Robert Bowen, University of California, Berkeley. Define f_k by $f_k(a_n)=b_n$ for $n\leqq k$ and $f(x)=0$ otherwise. Clearly the f_k are of class 1; as $f=\lim f_k$, f is of class $\leqq 2$. Suppose $\lim b_n=0$. Then

$$\sup |f(x) - f_k(x)| = \sup \{ |b_n| : n > k\} \to 0 \quad \text{as } k \to \infty;$$

$f_k\to f$ uniformly, and thus f is of class 1. On the other hand (ii) does not imply

(i) for if $a_n = n$, then f has only countably many discontinuities and is of class 1, but b_n may be chosen so that b_n does not approach zero.

Also solved by Béla Bollobus (Hungary), M. D. Mavinkurve (India), Necdet Ucoluk, and the proposer.

Ascending Chain Condition in $\mathrm{Hom}_R(M, M)$

5309 [1965, 674]. *Proposed by E. H. Feller, University of Wisconsin, Milwaukee*

Let M be a finitely generated unitary right R module, where R satisfies the ascending chain condition for right ideals. Find an example to show that the collection of all R-homomorphisms from M to M does not satisfy the ascending chain condition for right ideals. The elements of $\mathrm{Hom}_R(M, M)$ are written on the right.

Solution by A. G. Heinicke, University of British Columbia. Let R be a ring with an identity, and set $M = R$, considered as a right R-module. The mapping of $\mathrm{Hom}_R(R, R)$ into R, where $f \in \mathrm{Hom}_R(R, R)$ maps to $(1)f$, is easily seen to be an anti-isomorphism onto the ring R. Thus we may use any ring which has an identity and which satisfies the ascending chain condition for right ideals, but not for left ideals. For one example see Cartan and Eilenberg, *Homological Algebra*, p. 16.

The Number of Permutations Written with Prescribed Cycles

5310 [1965, 794]. *Proposed by S. W. Golomb, University of Southern California*

Represent each of the $n!$ permutations on n letters as a product of disjoint cycles, and let T_n^k be the number of such permutations for which the longest cycle has at most k letters. Show that

$$T_{n+1}^k = \sum_{j=0}^{k-1} (n)_j T_{n-j}^k,$$

where $(n)_j = n!/(n-j)!$.

I. *Solution by C. A. Church, Jr., West Virginia University.* It is immediate that $T_n^1 = T_1^n = 1$ and $T_n^{n+m} = T_n^n$ for $m \geqq 0$. We take the permutations from among the first $n+1$ natural numbers. For $k \geqq 2$ the occurrences of $n+1$ fall into the k mutually exclusive cases: $n+1$ occurs in a cycle of length $j+1$, $j = 0, 1, \cdots, k-1$. The number of distinct cycles of length $j+1$ in which $n+1$ occurs is $(n)_j$ and the complementary cycles are precisely T_{n-j}^k. Thus each case accounts for $(n)_j T_{n-j}^k$ selections. Summation over j from 0 to $k-1$ gives the result.

The solution is also contained in equation (5), p. 70 of J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley (1958).

II. *Solution by L. Carlitz, Duke University.* The number of permutations with r_1 unit cycles, r_2 2-cycles, $\cdots$, r_k k-cycles is equal to

$$\frac{n!}{1^{r_1}r_1!2^{r_2}r_2!\cdots k^{r_k}r_k!}, \quad (r_1 + 2r_2 + \cdots + kr_k = n).$$

(See Riordan, loc. cit., p. 67.) Thus

$$T_n^k = \sum \frac{n!}{1^{r_1}r_1!2^{r_2}r_2!\cdots k^{r_k}r_k!}$$

summed over all $r_1, r_2, \cdots, r_k$ such that $r_1+2r_2+\cdots+kr_k=n$. It follows that

$$\sum_{n=0}^{\infty} T_n^k \frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2} + \cdots + \frac{x^k}{k}\right),$$

so that

$$\sum_{n=0}^{\infty} T_{n+1}^k \frac{x^n}{n!} = (1 + x + \cdots + x^{k-1}) \sum_{n=0}^{\infty} T_n^k \frac{x^n}{n!}.$$

Equating coefficients, we get the proposed formula.

Also solved by I. N. Baker (England), Harley Flanders (England), S. Gobel (Netherlands), D. A. Hejhal, M. G. Greening (Australia), Edward Hook, C. C. Lindner, P. A. Morris (Jamaica), Ralph Parris (Jamaica), and the proposer.

The Integral of a Normalized Polynomial with Real Roots

5311 [1965, 794]. *Proposed by D. Ž. Djoković, University of Belgrade, Yugoslavia*

Let $x_1 < x_2 < \cdots < x_n$ be real numbers and

$$f(x; x_1, \cdots, x_n) = (x - x_1)(x - x_2)\cdots(x - x_n),$$

$$M = \max_{x_1 < x < x_n} |f(x; x_1, \cdots, x_n)|,$$

$$\phi(x_1, x_2, \cdots, x_n) = \frac{1}{M}\int_{x_1}^{x_n} f(x; x_1, \cdots, x_n)dx.$$

Prove (or disprove) the inequality $(-1)^k(\partial\phi/\partial x_k) > 0$.

Editorial Note. Dennis A. Hejhal, a student at Lane Technical High School, Chicago, shows that the inequality is untrue even for $n=2$. For then $f(x; x_1, x_2) = (x-x_1)(x-x_2)$, $M=(x_1-x_2)^2/4$, and ϕ is easily computed. We have $\phi(x_1, x_2) = 2(x_1-x_2)/3$, so that $\partial\phi/\partial x_1 > 0$ and $\partial\phi/\partial x_2 < 0$.

No other solutions or comment have been received.

From a reexamination of the original proposal there is reason to think that the inequality should have been stated as

$$(-1)^{n+1-k}(\partial\phi/\partial x_k) > 0.$$

A Vandermonde Operator

5312 [1965, 795]. *Proposed by D. S. Mitrinović, University of Belgrade, Yugoslavia*

Prove (or disprove) the equation

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \cdots & \dfrac{\partial}{\partial x_n} \\ \cdot & \cdot & \cdot & \cdot \\ \dfrac{\partial^{n-1}}{\partial x_1^{n-1}} & \dfrac{\partial^{n-1}}{\partial x_2^{n-1}} & \cdots & \dfrac{\partial^{n-1}}{\partial x_n^{n-1}} \end{vmatrix} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \cdot & \cdot & \cdot & \cdot \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = 1!2! \cdots n!.$$

Solution by D. G. Kabe, Northern Michigan University. The result is correct. Let us put $\sum(\pm)x_1^{r_1}x_2^{r_2} \cdots x_n^{r_n}$ for the (right-hand) determinant and $\sum(\pm)\partial_1^{r_1}\partial_2^{r_2} \cdots \partial_n^{r_n}$ for the (left-hand) Vandermonde operator, where $\partial_i = \partial/\partial x_i$, etc., and $(r_1, r_2, \cdots, r_n)$ is a permutation of $0, 1, \cdots, (n-1)$, and the ambiguous sign is determined by the usual procedure of expanding a determinant. Each summation contains $n!$ terms.

We see that there are exactly $n!$ terms which do not vanish in the procedure of differentiation, i.e.

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \partial_1 & \partial_2 & \cdots & \partial_n \\ \cdot & \cdot & \cdot & \cdot \\ \partial_1^{n-1} & \partial_2^{n-1} & \cdots & \partial_n^{n-1} \end{vmatrix} \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \cdot & \cdot & \cdot & \cdot \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}$$

$$= \sum \partial_1^{r_1}\partial_2^{r_2} \cdots \partial_n^{r_n} x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n},$$

where the summation is carried over $n!$ permutations of the integers $r_1, r_2, \cdots, r_n$. The equation follows because each term in the expansion is $0!, 1!, 2!, \cdots, (n-1)!$ in some order.

Also solved by R. A. Adams, W. T. Bailey, P. N. Bajaj, V. K. Batra (India), L. Carlitz, Yu Chang & Sidney Spital, G. W. Dinolt, L. O. Ferguson, Harley Flanders, B. L. Foster, Philip Fung, Sylvan Greene, Stephen Hoffman, D. A. Hejhal, R. R. Janić (Yugoslavia), R. F. Jessup, D. C. Kay, P. G. Kirmser, P. V. O'Neil, E. A. Parent, C. B. A. Peck, J. R. Purdy, Kenneth Rogers, J. T. Rosenbaum, J. J. Schäffer (Uruguay), Donna J. Seaman, M. J. Sheridan, Kao-Hwa Sze, J. I. Thornby, C. A. Tsonis, C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), A. Weinmann (England), and P. H. Young.

Editorial Note. Several solvers interpreted the problem differently and proceeded to effect a multiplication of the determinants. Such procedure is considered invalid because the result is dependent on the manner in which the multiplication is effected.

On Circumcenters and Orthocenters

5313 [1965, 795]. *Proposed by J. M. Quoniam, Saint-Etienne (Loire), France*

If ABC is a triangle with circumcenter O, orthocenter H, circumradius R, and area Δ, prove that

$$\cos 2A + \cos 2B + \cos 2C = \frac{OH^2}{2R^2} - \frac{3}{2}, \tag{1}$$

$$\sin 2A + \sin 2B + \sin 2C = 2\Delta/R^2. \tag{2}$$

Solution by William W. Willson, Cheltenham Grammar School, England. Write $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}$, $\overrightarrow{OC} = \mathbf{c}$. Then if $\overrightarrow{OX} = \mathbf{x} = \mathbf{a}+\mathbf{b}+\mathbf{c}$, we have $(\mathbf{x}-\mathbf{a})\cdot(\mathbf{b}-\mathbf{c}) = \mathbf{b}^2-\mathbf{c}^2 = 0$, so that AX is perpendicular to BC. Thus, by symmetry, X is the orthocenter, H.

Now $(\mathbf{a}+\mathbf{b}+\mathbf{c})^2 = \sum \mathbf{a}^2 + \sum 2\mathbf{b}\cdot\mathbf{c}$. That is (since $\sphericalangle BOC = 2\sphericalangle A$)

$$OH^2 = 3R^2 + 2R^2(\cos 2A + \cos 2B + \cos 2C).$$

Again $(\mathbf{c}-\mathbf{b})\times(\mathbf{a}-\mathbf{b}) = \mathbf{b}\times\mathbf{c}+\mathbf{c}\times\mathbf{a}+\mathbf{a}\times\mathbf{b}$, whence $2\Delta = R^2(\sin 2A + \sin 2B + \sin 2C)$.

Also solved by A. N. Aheart, D. P. Ambrose (Basutoland), P. N. Bajaj, V. K. Batra (India), W. J. Blundon, Robert Breusch, R. Buckley (England), L. Carlitz, P. Carragher, Ragnar Dybvik (Norway), Mrs. A. C. Garstang, M. G. Greening (Australia), Louise S. Grinstein, W. E. Hoff, Stephen Hoffman, Mrs. M. R. Iyer (India), R. R. Janić (Yugoslavia), D. G. Kabe, F. Leuenberger (Switzerland), Norman Miller, Q. G. Mohammad & M. L. Tikoo (India), G. L. N. Rao (India), Simeon Reich (Israel), H. A. Robinson, H. Simpson (England), R. Sivaramakrishnan (India), Sidney Spital, Al Somayajulu, Kao-Hwa Sze, P. D. Thomas, C. H. Tsonis, A. Vandeghen (Belgium), C. Van de Vyle (Belgium), Simon Vatriquant (Belgium), C. S. Venkataraman (India), and the proposer.

Iterates of a Hölder Function Transformation

5315 [1965, 795]. *Proposed by G. A. Heuer and R. G. Lee, Concordia College, Moorhead, Minn.*

If f is a function integrable on $[0, 1]$, let $Af = A^1 f$ be the function defined on $(0, 1]$ by

$$Af(x) = (1/x)\int_0^x f(t)dt; \qquad Af(0) = 0.$$

Let $A^{n+1}f = A(A^n f)$ for $n = 1, 2, \cdots$; obviously if f has a limit at 0, so does Af.

(1) Find a function f_0 bounded and continuous on $(0, 1]$, such that f_0 does not have a limit at 0 but Af_0 does.

(2) Find a function f_1, bounded and continuous on $(0, 1]$, such that Af_1 does not have a limit at 0.

(3) Does there exist, for each positive integer n, a function f_n, continuous and bounded on $(0, 1]$, such that $A^n f$ has no limit at 0?

(4) Is there a continuous bounded function f on $(0, 1]$ such that for every positive integer n, $A^n f$ fails to have a limit at 0?

I. *Solution by J. T. Rosenbaum, State University of New York at Buffalo.* Part 1. Let $f_0(x)=e^{i/x}(0<x\leqq 1)$. f_0 has no limit at 0 but

$$(Af_0)(x) = ixe^{i/x} - (2i/x)\int_0^x te^{i/t}dt \to 0 \quad \text{as} \quad x\to 0.$$

Parts 2–4. It suffices to show that part 4 should be answered in the affirmative. This follows immediately from the fact that $f(x)=\exp(i \log x)$ $(0<x\leqq 1)$ is an eigenfunction of A corresponding to the eigenvalue $1/(1+i)$.

II. *Remark by Roy O. Davies, The University, Leicester, England.* Significantly, we are not asked for a bounded continuous f on $(0, 1]$ such that for some positive integer n, A^nf does not have a limit at 0 but $A^{n+1}f$ does. Nonexistence of any such f follows from the observation that if f is bounded and continuous, then for $n\geqq 1$, $xA^nf(x)$ satisfies a Lipschitz condition

$$|xA^nf(x) - yA^nf(y)| \leqq k\cdot |x-y|.$$

We may then apply the following Tauberian

THEOREM. *If* $|x\phi(x)-y\phi(y)|\leqq k\cdot|x-y|$ *for* $0<x\leqq y\leqq 1$ *and* $A\phi(x)\to l$ *as* $x\to 0^+$ *then* $\phi(x)\to l$ *as* $x\to 0^+$.

Proof. We may take $l=0$. Suppose that $\limsup \phi(x)>\delta>0$. Then there exist arbitrarily small $x'>0$ for which $\phi(x')\geqq\delta$, that is $x'\phi(x')\geqq x'\delta$. It can easily be deduced from the Lipschitz condition that $x\phi(x)\geqq x(\frac{1}{4}\delta)$ for $x'\leqq x\leqq x''$ $\equiv\{(k+\delta)/(k+\frac{1}{2}\delta)\}x'$, and therefore $R(x')\equiv[1/(x''-x')]\int_{x'}^{x''}\phi(x)dx\geqq\frac{1}{4}\delta$, contradicting the fact that $R(x')=[x''/(x''-x')]A\phi(x'')-[x'/(x''-x')]A\phi(x')\to 0$ as $x'\to 0^+$.

Also solved by D. R. Anderson, I. N. Baker (England), Dennis A. Hejhal, J. G. Mauldon (England), A. Meir & A. Sharma, Donald Schaefer, and Sidney Spital.

Baker also notes that $A^nf(x)\to l$, $n>1$ implies that $Af(x)\to l$ if $f(x)$ is bounded and Lebesgue integrable. See Hardy & Littlewood, *Solution of the Cesaro summability problem for power series and Fourier series*, Math. Zeit., 19 (1923) 67–96.

Characteristic Roots of Products of Semi-definite Hermitian Operators

5316 [1965, 796]. *Proposed by J. T. Fleck and Carl Evans, Cornell Aeronautics Laboratory, Buffalo, N. Y.*

Show that if A and B are nonnegative definite Hermitian matrices of order n, the characteristic roots of AB are real and nonnegative.

I. *Solution by L. N. Howard, Massachusetts Institute of Technology.* A and B both have nonnegative definite square roots. Let x be any eigenvector of AB: $ABx=\lambda x$. If $B^{1/2}x=0$, then $\lambda=0$. Otherwise

$$\lambda|B^{1/2}x|^2 = \lambda x^*Bx = x^*BABx = |A^{1/2}B^{1/2}x|^2.$$

Thus $\lambda\geqq 0$.

II. *Solution by D. Topping, University of Washington.* We prove, more generally, that if A and B are positive semidefinite bounded Hermitian operators on a Hilbert space, then the spectrum of AB is real and nonnegative. First note that in any ring with identity, existence of $(1-xy)^{-1}$ implies existence of $(1-yx)^{-1}$; in fact, $(1-yx)^{-1}=1+y(1-xy)^{-1}x$. This implies that $\mathrm{sp}_0(AB)=\mathrm{sp}_0(BA)$, where $\mathrm{sp}_0(A)$ denotes the spectrum of A with zero adjoined. (In the finite-dimensional case, 0 is an eigenvalue of both AB and BA, or of neither.) Hence $\mathrm{sp}_0(AB)=\mathrm{sp}_0(A^{1/2}BA^{1/2})$ and since $A^{1/2}BA^{1/2}$ is positive semidefinite, the result follows.

III. *Solution by Sidney Spital, California State Polytechnic College, Pomona.* We form the positive definite Hermitian matrices $A(\epsilon)$ and $B(\epsilon)$ respectively by adding any positive ϵ to each diagonal element of A and B. The inverse $A^{-1}(\epsilon)$ therefore exists and is also Hermitian positive definite. Under these circumstances it is known that: (a) the eigenvalues, $\{\lambda_i(\epsilon)\}$, of $A(\epsilon)B(\epsilon)$ are the same as those obtained from the zeros of $\det[B(\epsilon)-xA^{-1}(\epsilon)]$; and (b) the latter are all real and positive. [See F. B. Hildebrand, *Methods of Applied Mathematics*, Prentice-Hall, (1952) pp. 74–79.] Hence all $\lambda_i(\epsilon)$ are real and positive. Since characteristic roots are continuous functions of their matrix elements and since the elements of $A(\epsilon)B(\epsilon)$ are (at most quadratic) polynomials in ϵ, it follows that $\lambda_i(0)$, the eigenvalues of AB, satisfy $\lambda_i(0)=\lim_{\epsilon\to 0}\lambda_i(\epsilon)\geqq 0$ (and of course are real).

IV. *Note by J. Wesley, Union College, Barbourville, Ky.* The theorem in problem 5316 was given in Carlson's solution of E 1677 [1965, 189].

The case of positive definite matrices, is Theorem 11.15, p. 92, Faddeev and Faddeeva, *Computational Methods of Linear Algebra*, Freeman, San Francisco, 1963.

Also solved by R. A. Adams, C. G. Cullen, Roy O. Davies (England), Harley Flanders (England), J. A. Hejhal, D. G. Kabe, Marvin Marcus & Henryk Minc, O. J. Nikolai, R. F. Rinehart, Murray Schechter, H. Simpson (England), R. C. Thompson, J. F. Watters & A. Weinmann (England), J. P. Williams, and the proposers.

A Theorem in algebra with socio*logical* implications. Kernel Homomorphism decided to get his group together and form Equivalence Classes, but a diversionist, the Natural Isomorphism, induced the Equivalence Classes to march off to the Homomorphic Image.

J. E. SPARKS, University of Texas

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: KENNETH O. MAY, University of Toronto and E. P. VANCE, Oberlin College

Materials intended for review should be sent to Prof. May, Amer. Math. Monthly, Dept. of Math., University of Toronto, Toronto 5, Ont., Canada.

Knot Groups. By L. P. Neuwirth. Princeton University Press, Princeton, N. J., 1965. 111 pp. $3.50.

This excellent research monograph is Number 56 in the Annals of Mathematics Studies. It is concerned with the fundamental group of the complement of tame polygonal knots in the 3-sphere. Topics discussed include a combinatorial covering space theory for 3-manifolds, subgroups and the Alexander matrix, representations, automorphisms, and the problem of characterizing such groups by their algebraic properties. One chapter is devoted to the construction of a group of groups which is given a categorical setting in an appendix by S. Eilenberg.

The book is clearly and lucidly written. In particular, the geometric expositions are done very pleasantly, with good figures and a few excellent pictures of knots. Many new results are proved here, while for previously known theorems the proofs have frequently been omitted. The last chapter is a collection of research problems.

R. P. JERRARD, University of Illinois

Integration of Equations of Parabolic Type by the Method of Nets. By V. K. Saul'yev. Pergamon, New York, 1964. Translated from the Russian edition published by Fizmatgiz, Moscow, 1960 by G. J. Tee. Translation edited by K. L. Stewart. xvii+346 pp. $12.00.

This work is divided into two parts. Part I is devoted to the construction of net equations (difference equations) approximating a parabolic equation. The emphasis is mainly on the one-dimensional heat equation (113 pages). There is a very clear and thorough study of the stability and the accuracy of many methods, both classical and new. There is also a table summarizing the advantages and shortcomings of each method. Methods for solving the heat equation in higher dimensions are also examined carefully (46 pages). Finally, 33 pages are devoted to more general parabolic equations.

Part II of the book (the final 1/3) is devoted to practical numerical solution of implicit net equations. This is essentially an elliptic problem because the time is held constant. Again the discussion is clear and thorough.

The book is written for a wide audience. The author assumes that the reader knows linear algebra and has some familiarity with the heat equation. The translation is smooth and accurate, and the translator's notes are helpful. The publisher is to be commended for offering this worthwhile book.

G. W. HEDSTROM, The University of Michigan

Ideale Ränder Riemannscher Flächen. By C. Constantinescu and A. Cornea. (*Ergebnisse der Mathematik und ihrer Grenzgebiete*), Springer Verlag, Berlin-Göttingen-Heidelberg, 1963. viii+244 pp. DM 68.00.

This book gives a unified and self-contained treatment of the theory, which has developed over the last fifteen years, of harmonic functions on open Riemann surfaces, especially that of the classes HB and HD. It begins by considering subharmonic functions, the lattice of positive harmonic functions, the Dirichlet problem (à la Perron and Wiener), potentials of mass distributions with the Green's function as kernel, and energy and capacity in the Frostman fashion.

From here on the book centers around the concept of compactification: a compactification of an open Riemann surface R is a compact Hausdorff space R^* which contains R as a dense open subset. The difference $\Delta = R^* \sim R$ is called the ideal boundary for the compactification. One wants to construct compactifications for R so that various classes of functions on R admit of extensions to R^* in some sense and so that it is meaningful to consider boundary value problems (say for harmonic functions) with data prescribed on Δ.

If Q is a class of continuous functions on R, then a Q-compactification $R^* = R_Q^*$ of R is a compactification such that each function in Q can be extended to a continuous function on R^* and these extended functions separate the points of Δ. Using standard structure theory, one can show easily that for each Q there is a unique Q-compactification. If we take for Q all bounded continuous functions, then R_Q^* is the Stone-Čech compactification of R. If Q consists of all differentiable functions on R such that df has compact support (i.e., $df \equiv 0$ outside a compact set), then R_Q^* is the Kerékjártó-Stoilow compactification of R. These two compactifications depend only on the topology of R and are the largest and smallest compactifications which have certain natural separation properties. Useful compactifications for potential theory on R are obtained by making suitable choices of Q between these extremes.

The authors call a function f on a hyperbolic surface harmonizable if the supremum of those subharmonic functions which are less than f except on a compact set is the same as the infimum of the class of superharmonic functions which are greater than f except on a compact set. A continuous function f on a Riemann surface R is called a Wiener function if for each hyperbolic subdomain G of R the restriction of f to G is harmonizable and $|f|$ has a superharmonic majorant on G. A continuous function f is called a Dirichlet function if it has a differential (in the weak sense) which is square integrable. The authors call the Q-compactification of R the Wiener compactification if Q is the class of continuous Wiener functions on R and the Royden compactification if Q is the class of all Dirichlet functions on R. Much of the first half of the book is devoted to the construction and properties of these compactifications.

The former of these compactifications is intimately connected with the theory of the class HB of bounded harmonic functions on R, and the latter

with the class HD of harmonic functions with a finite Dirichlet integral. Boundary value problems and connections with potential theory are explored at some length.

By their very definition these compactifications are well suited to the study of the structure of the classes HB and HD. Unfortunately, they are fairly large compactifications, and one often wants smaller ones. For example, if R is the interior of the unit circle, then the Wiener compactification does not just add the circumference but instead adds the Stone space of the bounded measurable functions on the circumference. This "defect" leads to the consideration of two other ideal boundaries. The first is the Martin boundary obtained by adding points corresponding to the minimal positive harmonic functions. In the case that R is the interior of the unit circle, then the Martin boundary is the unit circumference. The idea of this boundary goes back to the work of Martin in 1941 who used such a construction to regularize the Dirichlet problem for regions in three-dimensional space. It has received much attention by the modern French school of potential theory, and this book gives a thorough discussion of it in so far as Riemann surfaces are concerned. The Martin boundary is related to the space of positive harmonic functions, and of bounded harmonic functions. The Kuramochi boundary plays a similar role for the space HD of harmonic functions with a finite Dirichlet integral. Much of the last half of the book is devoted to the properties and application of the Martin and Kuramochi boundaries.

HALSEY ROYDEN, Stanford University

Problems in the Sense of Riemann and Klein. By Josip Plemelj. Interscience, New York, 1964. vii+175 pp. $8.00.

This is a book about one aspect of classical analysis written by a man who played an important part in its development. The problem of Klein is concerned with the mapping properties of the ratio of two linearly independent solutions of a Fuchsian differential equation. That of Riemann seeks the determination of n functions analytic in a given domain with limit at the interior side of the closed boundary equal to a linear combination of the exterior limit values of n given functions analytic in the exterior. The solution of these problems leads to certain results in the theory of integral equations.

Part I considers the theory of ordinary linear differential equations with analytic coefficients. There is a detailed investigation of the hypergeometric equation and of the second order Fuchsian equation. Klein's theorems about the behavior of solutions and a short chapter on oscillation conclude this part.

Part II treats integral equations, in particular the Fredholm equation. The book concludes with a presentation of the Riemann problem in which the functions concerned have a pre-assigned monodromy group.

There are no exercises.

STEPHEN HOFFMAN, Trinity College

Algèbre, M. G. P. et Spéciales A. By Michel Queysanne. Collection U, Série "Mathématiques." Directed by André Revuz. Librairie Armand Colin, Paris, 1964. 608 pp. 51,00 F.

This is the first volume of a projected three volume work on mathematics; later volumes, by other authors, will treat analysis and geometry. It is a big book, intended to cover thoroughly the algebra required in two programs leading to certification of mathematics and science students in France. Up-to-date in style, it will be welcomed for its clear presentations of many of the topics which have become traditional in "modern algebra" and for its excellent sequences of exercises.

The first three chapters are on sets, mappings and relations, the natural numbers, and composition laws. Here the author presents the basic terminology, occasionally using simple commutativity diagrams. Properties of finite sets and elementary facts of combinatory analysis are developed in the language of injections, surjections, and bijections. In his choice of terms and notations the author follows Bourbaki: he speaks of internal and external composition laws, neutral elements, and stable subsets; he stresses the notion of an equivalence relation compatible with given operations; he includes remarks on the notion of a structure.

The next three chapters deal with groups, rings and fields, and the complex numbers. A look at the chapter on groups may give some idea of the general depth of the treatment. It includes material on subgroups, Lagrange's theorem, quotient groups, the canonical decomposition of a homomorphism, direct sums, generators, cyclic groups, permutation groups, and Cayley's theorem. It does not include material on semigroups, on the isomorphism theorems or the Jordan-Hölder-Schreier theorem, on the Sylow theorems, on free groups, or on exact sequences.

Chapters 7, 8, 9 and 10 discuss vector spaces, matrices, determinants, and linear equations in that order. Modules are not treated, except in two exercises. The determinant is defined as a certain n-linear alternating form; exterior algebra and tensor product are not mentioned.

Polynomials, rational fractions, and algebraic equations are the subjects of Chapters 11, 12 and 13. Along with a study of the polynomial rings $K[X]$ and $K[X_1, \cdots, X_m]$, where K is a commutative field, these chapters contain sections on symmetric polynomials, on partial fractions, and on elimination, resultants and discriminants.

The final two chapters concern characteristic values and vectors of an endomorphism, the reduction of matrices to diagonal and triangular forms, and symmetric bilinear and hermitian forms.

A most valuable feature of this text is its collection of 559 exercises, which enable the reader to test and to strengthen his comprehension; some of these introduce new concepts or advance the theory beyond what is explained in the text. Considering these exercises, as well as the uniformly high quality of the

exposition, one does not hesitate to recommend speedy translation of the present book into English. We may also look forward to seeing the other volumes of this series.

G. N. RANEY, University of Connecticut

Symbolic Logic and Language, a Programmed Text. By James Dickoff and Patricia James. McGraw-Hill, New York, 1965. iv+390 pp. $6.95.

This book is mistitled. Its 390 pages are devoted solely to the propositional calculus. The authors promise us a sequel, *Symbolic Logic and Systems*, dealing with the predicate calculus and the study of formal systems.

By numerous examples the authors try to explain the meaning of *logically valid argument.* Then they develop a natural-deduction system F which is a classical form of an intuitionistic system of Fitch. The completeness of system F is mentioned, but not proved. Indeed, it could hardly be proved, since the authors never give a precise definition of *logically valid.* It is true that in an appendix they sketch the ideas of truth tables and tautologies, but they never indicate that this approach yields the most natural definition of *logically valid argument.* A Hilbert-type system (due to Lukasiewicz) is also introduced in the appendix and it is stated that this system, the approach via truth tables, and system F give essentially the same results.

No interesting metamathematical properties of the system F are studied. In fact, if we except proofs that certain specific formulas are provable in F, there is not one proof in the whole book. The shallowness of the text seems to be a consequence of its method, that of a learning program. This works well when simple matters have to be taught so thoroughly that even a moron can understand. But it would be a miracle if someone were able to explain deep mathematical results using this technique. In addition, after the novelty has worn off, any adult above the moron level will become exasperated by the repetition of obvious points.

There are two minor details which need correction. First, in the body of the text the system F is provided with only five statement letters. This situation is corrected in Appendix A, where an unlimited number of letters is allowed, but the authors characterize this as a "trivial extension of F." The second point concerns the translation of "A unless B" as "A if and only if not-B." The inaccuracy of this translation is revealed by one of the authors' own examples: "Napoleon was defeated at Waterloo unless history deceives us." Surely history could deceive us even though Napleon was defeated at Waterloo. A more accurate translation of "A unless B" seems to be "If not-A then B" (or, equivalently, "If not-B then A").

The good points of this book should not be overlooked. It has very thorough drill on translation from English into the symbolism of F and on the invention of proofs in the system F.

ELLIOTT MENDELSON, Queens College

Algebra of Matrices. By Malcolm F. Smiley. Allyn and Bacon, Boston, 1965. xii+258 pp. $8.75.

This book is a well organized, neatly executed, rapidly paced text, the outgrowth of a one semester course taught for a number of years by the author at the State University of Iowa and the University of California, Riverside.

The scope of the text is indicated by the chapter headings: Rings and Matric Rings, Determinants, Polynomials, Matric Polynomials and Functions of Matrices, Vector Spaces, Rank and Linear Equations, Finite Games, Linear Transformations and Matrices, Similarity of Matrices, Euclidean Spaces, Unitary Spaces, Quadratic Forms and Witt's Theorem. An appendix on set theory is also included.

Each chapter begins with a short statement summarizing the content of the chapter and how it is related to the overall development of the subject. Exercises, of which there are over 300 of various degrees of difficulty, are grouped at the ends of chapters. No answers to exercises appear in the text. There are ample illustrative examples.

The author works in a more general framework than that found in most introductory texts on the subject. Where possible, scalars are elements of a division ring and occasionally this restriction is weakened or modified.

The author assumes "that the reader has already had a good course in calculus." Knowledge of calculus is not really necessary (the few sections in which calculus is used may be omitted without any critical loss), but the level of mathematical maturity reached by successfully completing a good calculus course is necessary to read the text profitably. For the audience for whom the author writes, this book should prove to be an excellent text.

E. A. MAIER, University of Oregon

Group Theory. By W. R. Scott. Prentice-Hall, Englewood Cliffs, N. J., 1964. 479 pages. $10.50.

A wide variety of topics in many important areas of contemporary group theory is offered. The basic facts for each topic are stated concisely and proved in full, and deeper results are often given. Despite the author's essential clarity, reading is somewhat difficult in places because of the extensive use of special notation and symbolism. The reader must often translate masses of compressed data into English in order to see what is going on (page 111, for instance). Nevertheless, the effort is well worth it, and the book should prove useful to a wide variety of readers. The many exercises, ranging from the trivial to the challenging, help the text along and aid the student in filling in gaps in the theory. As a reference book its value is beyond question, especially for anyone who wishes to get into the field. As a text, it should be teachable, containing as it does more than enough material for a one-year course; but the instructor will have to prepare well in order to help his students through the high density regions.

FRANKLIN HAIMO, Washington University

Operations Research: Process and Strategy. By David S. Stoller. University of California Press, Berkeley, 1964. 159 pp. $5.00.

This book is divided into three parts; a short history and description of Operations Research, a reasonably extensive treatment of queuing theory, and a detailed presentation of zero-sum two person matrix games and solution methods. The chief prerequisite is a reasonable command of elementary probability and function concepts. On the whole, this book is well written, but the reader must be prepared to read through a good deal of symbolism and equations, even when only elementary concepts are involved. The only major disappointment is the treatment of the gradient-simplex method, which mentions none of the elegant bookkeeping schemes available for this algorithm and erroneously presents it as a process subject to indefinite repetition rather than as a process that can be made to converge nicely after a finite number of steps.

In summary, this book is too detailed to be recommended as a popular exposition and too narrow to be used as an exclusive text, but it does a good job on its limited objectives with the exception noted above.

R. E. STEARNS, General Electric Research and Development Center

Regular Figures. By L. Fejes Tóth. Macmillan, New York, 1964. xi+339 pp. $12.00.

This is in many respects a most charming book. It is in two parts, which are likely to be of interest to rather different types of readers. Part I deals with the geometry of the regular figures and their groups of symmetry, and is extremely lucid and well illustrated with figures. The first of its five chapters deals with plane crystallography, describing and illustrating the seventeen groups, as well as the point groups, and those that leave a line invariant. Here there are three extremely fine colored plates, of Egyptian, Arabic, and other ornamental designs. The next two chapters deal with similar questions on the sphere (including the 32 point groups of crystallography) and in the hyperbolic plane; and the other two with regular and semi-regular polyhedra, and with polytopes in higher space. The three dimensional figures are admirably illustrated in a series of anaglyphs (stereographic figures in two colors, with colored eyepieces provided) which are separate in a pocket at the end of the book.

Part II deals with all kinds of extremal problems, such as close packing and covering by spheres and other figures, maximum and minimum areas and volumes, and so forth, whose solution involves regular figures in some way; the maximal density of the hexagonal packing of circles, and the minimal area of the honeycomb, are typical. It is in five chapters, corresponding in scope to those of Part I, namely the Euclidean plane, the sphere, the hyperbolic plane, three dimensional space, and higher space. This part of the book seems to be, for the author, its main *raison d'être*, and it includes many original results, as well as much up-to-date work of others. For this reviewer it is much harder to read than the first part, but it will undoubtedly be of great interest and value to specialists in this field.

PATRICK DU VAL, University College, London

Finite Permutation Groups. By Helmut Wielandt. Academic Press, New York and London, 1964. x+114 pp. \$2.45 paper, \$5.50 cloth.

Anyone who wishes to learn about the explosive, new theory of finite groups would do well to start with this book as a background. After an easily readable introductory chapter, one goes on to a discussion of multiple and half transitivity. Chapter III is concerned with the structure of the permutation subgroup which fixes a given member of the permuted set. In Chapter IV, a group is looked upon as a permutation group on the set of one of its own subgroups, and subrings of the group ring are treated. Further, certain abstract finite groups which force the primitive permutation groups which extend them to be doubly transitive are likewise discussed here. The last chapter, on representation theory, acts as a unifying conclusion. Among the topics here are the centralizer ring, a vector space, the elements of which commute with the permutation matrices; matrix reduction; and primitive groups of degree $2p$ (p, a prime), where $2p \neq a^2+1$ for any integer a.

The text abounds with historical hints, and references are brought up to 1962. Those interested in the history of group theory in the United States will note that the contributions of W. A. Manning to the subject are documented on virtually every page. It would, perhaps, have been better, had the author expanded this slim volume somewhat so that proofs of some results, either totally suppressed or highly abbreviated here, could have been given in full. Nevertheless, no serious student of the subject should be without this book.

FRANKLIN HAIMO, Washington University

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor George Best, Phillips Academy, represented the Association at the Convocation in observance of the One Hundredth Anniversary of the founding of Dean Junior College on May 14, 1966.

Professor J. W. Cell, North Carolina State University, represented the Association at the Centennial Convocation and the inauguration of James E. Cheek as President of Shaw University on April 16, 1966.

Professor J. N. Eastham, Queensborough Community College, represented the Association at the inauguration of John S. Toll as President of the State University of New York at Stony Brook on April 16. He also represented the Association at the inauguration of George Ferguson Chambers as President of Nassau Community College on May 7.

Professor N. G. Gunderson, University of Rochester, represented the Association at

the inauguration of Albert W. Brown as President of State University College at Brockport on May 19, 1966.

Professor F. B. Jones, University of California, Riverside, represented the Association at the inauguration of John M. Pfau as President of California State College at San Bernardino on May 4, 1966.

Professor V. V. Latshaw, Lehigh University, represented the Association at the Engineering-Science Centennial Convocation at Lafayette College on April 15 and 16, 1966.

Professor E. P. Starke, Bloomfield College, represented the Association at the Bicentennial Convocation of Rutgers, The State University on September 22, 1966.

American University of Beirut: Mr. Robert Fraga, American University in Cairo, and Dr. David Singmaster, University of California at Berkeley, have been appointed Assistant Professors; Associate Professor Edward Batho, University of New Hampshire, has been appointed Visiting Professor; Associate Professor Amin Muwafi, Chairman of the Mathematics Department, will be on sabbatical leave as Visiting Professor at the University of Florida.

Lindenwood College: Mr. R. W. Murdock, Principia College, has been appointed Chairman of the Mathematics Department; Professor S. Louise Beasley will be on sabbatical leave for 1966–67 as a participant in a Scandinavian Seminar.

Assistant Professor G. H. Andrews, Oberlin College, has been promoted to Associate Professor.

Mr. C. R. Conniff, Wisconsin State University at La Crosse, has been granted a year's leave of absence to attend the University of Minnesota.

Mr. G. R. Costello, Aerospace Corporation, San Bernardino, California, has been promoted to Head of the Guidance Technology Department in the company's Technology Division.

Mr. E. B. Davis, Stanford University, has been appointed Assistant Professor at Lawrence University.

Assistant Professor D. L. Kreider, Dartmouth College, has been promoted to Associate Professor.

Professor Chia-Chiao Lin, Massachusetts Institute of Technology, has been named Institute Professor.

Professor K. O. May, on leave from Carleton College as visiting scholar at the University of California at Berkeley, has been appointed Professor at the University of Toronto and the Ontario College of Education.

Lt. Col. R. K. Moorhead, U. S. Air Force Academy, has been promoted from Associate Professor to Tenure Associate Professor.

Dr. Jordan Rosenbaum, State University of New York at Buffalo, has been appointed Assistant Professor at the University of Pittsburgh.

Mr. R. L. Shubert, University of Kansas, has been promoted to Assistant Professor.

Dr. J. L. Sieber, Shippensburg State College, has been promoted from Acting Chairman to Chairman of the Mathematics Department.

Dr. F. J. Weyl, formerly of the Office of Naval Research, has been appointed as a Special Assistant to the President of the National Academy of Sciences.

Assistant Professor John W. Wyman, Pasadena College, has been awarded an N.S.F. Science Faculty Fellowship at the University of Southern California.

Professor Emeritus G. A. Bingley, St. John's College, died on February 27, 1966. He was a charter member of the Association.

Mrs. Tempie R. Franklin, Arlington County Public Schools, Virginia, died on January 28, 1966. She was a member of the Association for six years.

Professor Alexander Wittenberg, York University, died in December 1965. He was a member of the Association for seven years.

THE UNIVERSITY OF TEXAS—FIFTH ANNUAL SYMPOSIUM ON BIOMATHEMATICS AND COMPUTER SCIENCE IN THE LIFE SCIENCES

The Division of Continuing Education of The University of Texas Graduate School of Biomedical Sciences at Houston is pleased to announce the Fifth Annual Symposium on Biomathematics and Computer Science in the Life Sciences, which will be held at the Shamrock Hilton Hotel in Houston, Texas, in March 1967. The theme of the symposium is MEDICAL USES OF MAN-MACHINE SYSTEMS. The topics for the sessions of the symposium are: "The Engineering and Mathematical Simulation of Artificial Organs," "Computer Data Editing Display and Graphic Systems," "Patient Monitoring," "Radiation Treatment Planning," "Research Laboratory Automation," "Medical Applications of Computer Hardware and Software Systems," "Training Program in Biostatistics, Biomathematics, Computer Sciences, and Bioengineering."

Those interested in submitting abstracts should contact the Office of the Dean, Division of Continuing Education, The University of Texas Graduate School of Biomedical Sciences at Houston, 102 Jesse Jones Library Bldg., Texas Medical Center, Houston, Texas 77025. Abstracts will not be reviewed for inclusion in the program after November 30, 1966.

EAI ANALOG COMPUTER EDUCATIONAL USERS GROUP

The EAI Analog Computer Educational Users Group has been formed in order to meet the needs of those desiring to use analog computers in the teaching of physics or in their research. This nonprofit organization issues a Newsletter and Application Notes on a regular basis. One of the recent Application Notes dealt with the compound pendulum and how it could be simulated on an analog computer. Further information about this Group and samples of its publications may be obtained by writing to EAI Analog Computer Educational Users Group, P.O. Box 90626, Los Angeles, Cal. 90009.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

NEW SECTIONAL GOVERNORS OF THE ASSOCIATION

The following have been elected Governors of the Association for the three year term July 1, 1966 to June 30, 1969, by a mail vote of the Association in the Sections indicated:

Allegheny Mountain	A. F. Strehler, Carnegie Institute of Technology
Indiana	G. N. Wollan, Purdue University
Kentucky	L. L. Scott, University of Louisville
Metropolitan New York	Abraham Schwartz, City University of New York
Nebraska	D. W. Miller, University of Nebraska
Northern California	D. W. Blakeslee, San Francisco State College
Oklahoma-Arkansas	R. B. Deal, Oklahoma State University
Rocky Mountain	F. M. Stein, Colorado State University
Wisconsin	C. J. Vanderlin, Jr., University of Wisconsin

The highest percentage of voters was 51% in the Nebraska Section, followed by the Oklahoma-Arkansas Section with 39%.

RAOUL HAILPERN, *Associate Secretary*

FEBRUARY MEETING OF THE LOUISIANA-MISSISSIPPI SECTION

The forty-third annual meeting of the Louisiana-Mississippi section of the MAA was held in Baton Rouge, Louisiana, on February 18–19, 1966 with Louisiana State University as host institution. There were 243 persons registered including 110 members of the Association.

The general session on Friday afternoon was presided over by Chairman S. R. Knox and included the welcome by Provost Cecil Taylor of Louisiana State University and the response. The regular Committees were appointed. Professor Gail Young presented a paper entitled "The General Mathematics Curriculum Proposals of CUPM."

The technical papers were presented in two concurrent sessions with Professor W. E. Koss, Vice-Chairman, and Professor Virginia Carlton, Vice-Chairman, presiding.

The following officers were selected for the coming year: Chairman, Robert C. Brown, Southeastern Louisiana College; Vice-Chairman for Mississippi, Donald A. King, Mississippi State College for Women; Vice-Chairman for Louisiana, Fred B. Wright, Tulane University; Secretary-Treasurer, Virginia Carlton, Centenary College.

The invited speaker for the meeting was Professor E. E. Moise, Harvard University. Dr. Moise spoke at the Friday evening banquet on the subject "Puzzles, Platonism, and Extroversion." His talk at the Saturday general session was entitled: "The Idea of a Concrete Deductive System."

Professor Peter Dembowski of the University of Wisconsin and the University of Frankfort presented an hour colloquium lecture; "Finite Geometries and Their Groups."

The following papers were presented:

1. *A knot projection has four small complementary domains*, by D. E. Penney, Louisiana State University, New Orleans.

2. *Remarks on the category of Boolean algebras and related topological categories*, by P. S. Schnare, Louisiana State University, New Orleans.

3. *Modern mathematics for parents*, by M. P. Berri, Tulane University; Dennis Nead, Louisiana State University, New Orleans; and Sr. Mary Robert Von Wolff, O. P., St. Mary's Dominican College.

During the spring semester of 1965, a noncredit seminar entitled *Modern Mathematics for Parents* was offered as a cooperative effort by members of mathematics departments of five Colleges and Universities in New Orleans. The seminar was intended to acquaint parents with some of the new topics being introduced into elementary and junior high school mathematics instruction.

Some details of the work of the committee and the conduct of the seminar are included in this paper. An evaluation of the effort and a summary of results as determined by a questionnaire distributed to participants are also included.

It is hoped that this will be of help to other groups interested in such a form of community service.

4. *A generalization of the Rayleigh probability density functions*, by A. M. Johnson, Louisiana Polytechnic Institute.

5. *Some results pertaining to inversion with respect to real central conics in the complex plane*, by N. A. Childress, University of Mississippi.

In the complex plane, the point P' can be defined to be the image of a point P in an inversion with respect to a real conic with center O if P' lies on line OP and $OP \cdot OP'$ equals the square of the distance from O to the point of intersection of line OP with the conic. The image of O is the ideal point, and conversely. This definition and its results can be used to define the polar of a point P with respect to a real central conic as a line passing through the inverse of P and having a certain slope. Some results of these definitions are examined.

6. *A conjecture on addition chains*, by C. T. Whyburn, Louisiana State University.

7. *A generalization of some trigonometric identities using differential equations*, by A. J. Zettl, Louisiana State University.

8. *A theorem on factorization of quaternions*, by Gordon Pall, Louisiana State University.

Let m and n be nonnegative integers not both even, and let $v=v_0+iv_1+jv_2+kv_3$ be an integral quaternion (i.e. with integral components $v_0, \cdots, v_3$) of norm mn. Then the number of factorizations $v=tu$, as a product of integral quaternions 5, u of respective norms m and n, is equal to $r_4(e)$, where $e=(v_0, v_1, v_2, v_3, m, n)$, and $r_4(e)=$ the number of representations of e as the sum of four squares.

9. *On the interchangeability of 2-links*, by W. C. Whitten, Jr., University of Southwestern Louisiana.

Let $L=K_1+K_2$ be an oriented link of two components tamely imbedded in S^3. In this note several interchangeability invariants of L are given, one of which is noted here. Assume that K_2 is a torus knot, and let $T=\{T_\alpha\}$ denote the collection of all unknotted tori on each of which K_2 lies. Let $C(K_1\cap T_\alpha)$ denote the cardinality of $K_1\cap T_\alpha$ for any T_α in T. Define $T_{K_1}(K_2)=\min \{C(K_1\cap T_\alpha) | T_\alpha \in T\}/2$. $T_{K_1}(K_2)$ shall be called the *torus linking number of* K_1 *with* K_2. Clearly, $T_{K_1}(K_2)=T_{K_2}(K_1)$, if L is interchangeable.

10. *The cluster points of the Riemann-Stieltjes sums*, by A. C. Pierce, Louisiana State University.

If f and g are real-valued functions on an interval $[a, b]$, the approximating sums for the Riemann-Stieltjes integral form a net on the partitions of $[a, b]$. Let $S_a^b fdg$ denote the cluster points of this net. Then if f_1+f_2 is g-integrable on $[a, b]$, then $S_a^b f_1 dg=\{\int_a^b (f_1+f_2)dg\}-S_a^b f_2 dg$. For any f, g defined on $[a, b]$, $S_a^b fdg=\{f(b)g(b)-f(a)g(a)\}-S_a^b gdf$. In neither of these equations may the subtracted terms be transposed to the left side.

11. *On unconditional bases in Banach spaces*, by J. R. Retherford, Louisiana State University

12. *Spaces of finite subsets of a topological space*, by R. M. Schori, Louisiana State University.

Z. L. LOFLIN, *Secretary-Treasurer*

FEBRUARY MEETING OF THE NORTHERN CALIFORNIA SECTION

The annual meeting of the Northern California Section of the MAA was held in Dwinelle Hall, University of California at Berkeley, on February 5, 1966. Program chairman was D. F. Coulter, Jr., of Hartnell College, Salinas. One hundred thirty persons attended the meeting, of whom one hundred were members of the Association.

A selection of books of interest to the Association was displayed during the morning. Three talks were given as follows:

1. *Mathematical foundations for national economic planning*, by David Gale, Brown University.

2. *The Heawood map-coloring conjecture*, by J. W. T. Youngs, University of California at Santa Cruz.

3. *Sets of constant width*, by G. D. Chakerian, University of California at Davis.

At a final session in the afternoon, R. H. McDowell addressed the section on the CUPM "General curriculum in mathematics for colleges." This was followed by a discussion; the chief contributors were M. Dreyfus of San Jose City College and L. Henkin of the University of California at Berkeley. The point was made that, at present, some courses cannot be taught in junior colleges until several universities classify the course substance as lower division material.

J. L. BRENNER, *Secretary*

MARCH MEETING OF THE KANSAS SECTION

The fifty-first annual meeting of the Kansas Section of the MAA was held at the University of Kansas, Lawrence, Kansas, on March 26, 1966, in conjunction with the annual meeting of the Kansas Association of Teachers of Mathematics. There were 371 persons registered including 83 members of the Section. Chairman Gilbert Ulmer presided at the morning and afternoon sessions.

The following officers were elected: Chairman, Jimmy Rice, Fort Hays State College, Hays; Vice-chairman, Sister Mary Paul Buser, Marymount College, Salina; Secretary-Treasurer, Helen Kriegsman, Kansas State College, Pittsburg.

A presentation on "The CUPM Report on a General Curriculum in Mathematics for Colleges" was given by G. S. Young. The following papers were presented:

1. *The close-to-convexity and univalence of an integral*, by W. M. Causey, University of Kansas.

Let S be the class of functions f regular and univalent in $|z|<1$ and normalized by $f(0)=0$, $f'(0)=1$. Define F by $F(z)=\int_0^z(f(t)/t)^\alpha dt$. Suppose $f\in S$ is close-to-convex with respect to g, $g(0)=0$, then F is close-to-convex with respect to G defined by $G(z)=\int_0^z(g(t)/t)^\alpha dt$ for any α, $0\leqq\alpha\leqq 1$. Also if $f\in S$ and α is any complex number $0\leqq|\alpha|\leqq(\sqrt{5}-2)/4$ then $F\in S$. For any α, $\frac{1}{2}<\alpha\leqq 1$, there exists a function $f\in S$ such that $F\notin S$.

2. *Some misconceptions in mathematics*, by R. D. Bechtel, Kansas State University.

Two examples of inconsistent arguments commonly found in materials for elementary school mathematics were presented. Related general questions were discussed.

3. *Preservation of pseudo-metrizability by quotient maps*, by C. J. Himmelberg, University of Kansas.

Let f be a function from X onto Y where X is a pseudo-metric space with pseudo-metric d. THEOREM: *If Y has the quotient topology relative to f, then Y is pseudo-metrizable iff for each open subset G of Y there exists a family $\{\epsilon(y)|y\in G\}$ of positive reals such that $N_{\epsilon(y)}(f^{-1}(y))\subset f^{-1}(G)$ for all $y\in G$, and $d(f^{-1}(y), f^{-1}(z))\geqq\epsilon_y-\epsilon_z$ for all $y, z\in G$.* To prove this theorem, the author obtains a proposition relating the quotient topology and the quotient uniformity on Y.

4. *Some results in the theory of near-rings*, by R. E. Williams, Kansas State University.

The definitions of D. W. Blackett [Proc. A.M.S., 4(1953) 772–785] are used with the exception of the radical of a near-ring, which is defined by G. Betsch [Math. Z., 78(1962) 86–90]. Blackett's result that a simple near-ring is a direct sum of minimal right ideals is compared to the corresponding result in ring theory. We give an example of a simple near-ring which is the direct sum of non-isomorphic minimal right ideals. We prove that if I is a minimal right ideal of a near-ring N with $J(I)\neq I$ and if I contains a nonzero right distributive element, then $I/J(I)$ is a near-field. We conjecture that $I=J(I)$ if N is a simple near-ring.

5. *Trigonometric sums and a method of Selberg*, by J. Chidambaraswamy and S. M. Shah, University of Kansas.

If the partial sums $S_n(x)$ of a trigonometric series $\sum\eta_n\cos(nx+c_n)$ are nonnegative over $[0, a]$ and if c_n's satisfy some conditions, then it is shown that $\sum\eta_n/n^c$ is convergent. Here c is a constant depending on $\{c_n\}$. A general theorem is also proved in which cos y is replaced by a bounded integrable function. If $\eta_n=1$ when n runs through a given sequence of values and 0 otherwise, an inequality is obtained for the minimum of the corresponding $S_n(x)$. These results extend recent results of Chowla. A technique for calculating the Fresnel-type integrals involved in inequalities for c_n is also given.

HELEN KRIEGSMAN, *Secretary-Treasurer*

MARCH MEETING OF THE SOUTHEASTERN SECTION

Emory University was host to the 45th annual meeting of the Southeastern Section of the MAA, March 25–26, 1966. Professor J. E. Maxfield, retiring Chairman of the

Section, and Professor Trevor Evans, Chairman of the Mathematics Department of Emory University, presided at the general sessions. The following one-hour invited addresses were scheduled for the general sessions: "Dilation Theory," by P. R. Halmos (Visiting Professor, University of Miami) and "Finite Fields," by Leonard Carlitz, (Duke University).

Total registration for this meeting was 340, including 72 non-members of the Association. The following officers were elected: Chairman, C. V. Aucoin (Clemson University); Vice-Chairman, A. D. Wallace (University of Florida). The invitation from Florida Presbyterian College to act as host to the 1967 meeting was reaffirmed and an invitation from East Carolina College, Greenville, N.C. to act as host to the 1968 meeting was accepted.

At the business meeting the following resolution was passed. "It is the sense of the SE Section that the Board of Governors continue to consider and explore the desirability and feasibility of playing a role in graduate mathematical education." There was also considerable discussion about the question of partitioning the Section, particularly with respect to the possible formation of a separate Florida Section of the MAA. The Secretary was relieved of the responsibility of conducting the poll voted at the previous meeting of the Section until the question of partition has been resolved.

The following charter members of the Association were recognized by name: Helen Barton, J. W. Lasley, Jr., Mrs. Mayme I. Logsdon, Eugenie M. Morenus, E. J. Moulton, A. R. Wapple, Fredrick Wood, K. B. Patterson.

The following contributed papers were presented:

1. *An expansion technique*, by O. R. Ainsworth, University of Alabama.

A purely formal method of deriving the Neumann expansion without recourse to the $O_n(t)$ polynomials is presented. The method is used to obtain the allied expansions $\sum b_n t^{n+1/2} J_{n+1/2}(at)$. The same method is used to obtain the expansion $\sum b_n L_n(at)$ without using the orthogonal integral $\int_0^\infty f(t)e^{-t}L_n(t)dt$.

2. *A note on a theorem of Simonsen-Dubreil*, by A. R. Bednarek and A. D. Wallace, University of Florida.

The purpose of this note is to expose a new proof and an "extension" of a theorem of W. Simonsen (Sur les Correspondances Multivoques entre deux Ensembles Abstraits, Acta Math., 81 (1949), 291–297) and (independently) M. L. Dubreil-Jacotin (Applications Multiformes et Relations d'équivalence, C. R. Acad. Sci., Paris 230 (1950), 906–908). Actual and potential applications of these theorems are considered.

3. *Enumeration of canonical sets by rank*, by J. V. Brawley, Jr., Clemson University.

Let $F=GF(q)$ denote a finite field of order q, and let F_n denote the $n\times n$ matrices over F. By a canonical set for F_n is meant a subset C of F_n with the property that each $A\in F_n$ is similar to one and only one member of C. In this paper a formula is obtained for the number $N(r:n, q)$ of matrices in C with rank r $(0\leqq r\leqq n)$.

4. *Higher degree recurrence relations*, Mrs. Barbara F. Chambers, University of Alabama.

A technique for obtaining recurrence relations among products of functions of the simple polynomial, hypergeometric type is developed. An assumed form of a higher degree recurrence relation is constructed with undetermined coefficients, which this technique then determines with great facility. The results can be used to evaluate integrals of products of functions, and this technique is applied to three typical polynomial sets; Legendre, Laguerre, and Hermite.

5. *A sequence of inclusion disks for a normal matrix*, by D. H. Clanton, Furman University.

This paper presents, for a normal matrix A of order n, an algorithm for the development of a sequence of normalized vectors $q^{(l)}$ $(l=0, 1, 2, \cdots)$. Associated with each vector $q^{(l)}$ is an inclusion disk whose center is the Rayleigh quotient, $q^H A q$ (the superscript H signifies the transposed conjugate), and whose radii form a decreasing sequence of vector norms. The sequence of norms

converge to zero and the sequence of Rayleigh quotients converge to a characteristic root of A with one exception. If a disk has two or more characteristic roots on its boundary, then the algorithm will not yield a new vector.

6. *Idempotents in group rings*, by D. B. Coleman, Vanderbilt University.

It is well known that if R is a ring of algebraic integers, and if G is a finite group, then the group ring RG has no idempotents except 0 and 1. It is not difficult to prove the following generalization of this result. THEOREM. *Let R be an integral domain (with 1) and let G be a group of order n. Then the group ring RG has a nontrivial idempotent if and only if some prime divisor of n is a unit in R.*

7. *Polynomials defined by a certain class of generating functions*, by Russell Cowan, University of Florida.

Several relations are obtained from the generating function defining the set of polynomials. Explicit expressions are obtained for the polynomials by solving a set of linear first order differential equations. The final form for a general polynomial is established by mathematical induction. The polynomials reduce to generalized Laguerre polynomials if the constants arising are chosen to satisfy certain simple conditions.

8. *A theorem on action by a continuum semigroup*, by Jane M. Day and A. D. Wallace, University of Florida.

Let T be a Bing (continuum semigroup) acting on X, a continuum with dense open half-line, W. Let C denote $x\backslash w$, the limit continuum in X, and suppose that cardinal $C>1$. Suppose also that T acts unitarily ($x\in Tx$ for each $x\in X$). Then either the action is trivial on some continuum properly containing C, or else C is the unique minimal T-ideal of X and X admits an abelian bing having the endpoint of X as its unit and C as minimal ideal and a group.

9. *Positive commutative semigroups on the plane*, by R. W. Farley, University of Tennessee.

A *real semigroup* is a topological semigroup containing a closed sub-semigroup R isomorphic to the multiplicative semigroup of real numbers, embedded so that 1 is an identity and 0 is a zero [J. G. Horne, Jr., "Real commutative semigroups on the plane," PJM, Vol. II (1961), pp. 981–987]. A *positive semigroup* is a topological semigroup containing a closed sub-semigroup N iseomorphic to the multiplicative semigroup of nonnegative real numbers, embedded so that 1 is an identity and 0 is a zero. Throughout, S is a positive commutative semigroup on the plane with no nilpotent elements. Necessary and sufficient conditions are given for S to be a real semigroup. If S is not a real semigroup, $H(1)$ must be connected. If S is the union of connected groups S has either two or an odd number of idempotents. There exist examples having any odd number of idempotents greater than or equal to five. The example with five idempotents is unique.

10. *Derivative relative of a vector about a system of reference oblique, non-rigid and moving*, by R. M. Fiterre, University of Alabama.

The formula gives geometrical interpretation to each of its four terms. Each term expresses the rate of change of the vector;

1st: About the system, when it is considered rigid and fixed.
2nd: When it is produced by the change of measures of the basis e_1, e_2, e_3, about the reciprocal vectors e^1, e^2, e^3.
3rd: By the instantaneous rotation of the system, when it is considered rigid.
4th: By the specific instantaneous rotation of each of the axes of $(o\ x\ y\ z)$.

11. *Newtonian analogues of the trigonometric and exponential functions*, by Tomlinson Fort, Emory University.

Let $x^{(n)}=x(x-1)(x-2)\cdots(x-n+1)$; $x^{(0)}=1$. Then let

$$\operatorname{sint} x = \frac{x^{(1)}}{1!} - \frac{x^{(3)}}{3!} + \cdots + (-1)^{n-1}\frac{x^{(2n-1)}}{(2n-1)!} + \cdots$$

$$\operatorname{cost} x = \frac{x^{(0)}}{0!} - \frac{x^{(2)}}{2!} + \cdots + (-1)^{n-1}\frac{x^{(2n-2)}}{(2n-2)!} + \cdots$$

$$\operatorname{expt} x = \frac{x^{(0)}}{0!} + \frac{x^{(1)}}{1!} + \cdots + \frac{x^{(n-1)}}{(n-1)!} + \cdots.$$

A detailed study is made of the functions defined by these series. It finally turns out that

$$\operatorname{sint} x = 2^{x/2}\sin\frac{\pi x}{4}, \qquad \operatorname{cost} x = 2^{x/2}\cos\frac{\pi x}{4}, \qquad \operatorname{expt} x = 2^{x}.$$

12. *The Cooperative College-School Science Project in Conecuh County, Alabama*, by W. L. Furman, Spring Hill College.

Schools of Conecuh County, Alabama, have been assisted in mathematics for four years through the programs of the Alabama Academy of Science. An NSF grant to Spring Hill College for a cooperative project with the county for 1966–67 provides for courses in the structure of the real number system, algebra, and geometry during the summer and academic year for twenty-five secondary school teachers. Most of the teachers in the county lack adequate preparation in mathematics. The project is designed to make a total impact on a rural county school system attempting to modernize its curriculum.

13. *A classification of partial differential equations via Riemannian geometry*, by B. A. Fusaro, University of South Florida.

The notion of associating a Riemannian space R^m with metric $ds^2 = g_{ij}dx^i dx^j$ and the operator $\Delta_2 = g^{ij}\nabla_i\nabla_j$ *via* $(g^{ij}) = (g_{ij})^{-1}$, and employing such concepts as characteristic cones goes back perhaps to Beltrami (1892). This was used extensively by Hadamard (1923) and M. Riesz (1949); more recently A. Douglis (1954), and others, have re-emphasized this geometric approach. The difficulties encountered in the study of corresponding equations with variable coefficients such as (*) $\Delta_2 u = 0$, often leads to a study of the special case of constant coefficients so that covariant differentiation reduces to ordinary differentiation, $\nabla_i = \partial_i$, and $g^{ij} = \delta^{ij}$, yielding (**) $\Delta u = 0$. The device of letting geometric subspaces supply a natural ordering to study problems like (*) does not seem to appear in the literature. With this point of view, (*) can be said to be in R^m, and (**) in E^m; intermediate cases of variable coefficients are given by the harmonic spaces H^m of Copson and Ruse (1939) and spaces of constant curvature.

This geometric classification of equations has been applied by the author (1965, Md.U. Ph.D. thesis). For example, a mean-value theorem of Åsgeirsson (cf. Courant-Hilbert, "Math-Physics," vol. II) hitherto valid only in E^m, has been extended to H^m.

14. *A generalization of the Laplace integral*, by M. O. Gonzalez, University of Alabama.

The author calls generalized Laplace integrals those of the form $\int_0^\infty e^{-st}F(s, t)dt$, where s is a complex parameter and t is real. Under certain assumptions concerning the function $F(s, t)$, a number of properties of the generalized Laplace integral are established which are similar to the well-known properties of the ordinary Laplace integral.

15. *Sets that are unions of two star-like sets*, by W. R. Hare, Jr. and J. W. Kenelly, Jr., Clemson University.

In this paper a condition is given which characterizes the closed subsets of a topological linear space which are the union of two star-like sets. This result answers a question posed implicitly in Valentine's *Convex Sets*.

16. *Concerning rational number bases*, by Reginald Mazeres, Tennessee Technological University.

This paper gives definitions and the arithmetic for rational bases. The author develops an algorithm for recognizing integers.

17. *The suggested impact of computer applications on the undergraduate mathematics curriculum —Report of Two Symposia*, by E. P. Miles, Jr., Florida State University.

Participants from industry, government, foundations and 45 collegiate mathematics departments met at Florida State University on March 14, 1966 to discuss curricular changes suggested by computer applications. Sponsored by ACM and SIAM, panelists Wm. F. Atchison, Wallace Givens, Herman Goldstine, Nat Macon, and F. J. Murray proposed and defended revised mathematics offerings appropriate for today's undergraduates. Then panelists Wm. Duren, John Hamblen, Russell Poor, and Leland Shanor commented on means of accomplishing the proposed curricular goals including implementing project TACT (Teaching Aids for Computer Techniques) through the Southern Center for Improved Instruction in Undergraduate Science and Mathematics being established by cooperative regional effort on the F.S.U. Campus.

18. *A plea for the terminology "Flexpoint,"* by Rev. T. F. Mulcrone, S.J., Spring Hill College.

The reasons which have induced the petroleum and chemical industries to discourage the use of the word "inflammable" and have brought about the legislation that explosive fluid carriers be labeled "Flammable" instead were applied to indicate the advantages of using some form of the root "flex" instead of "inflection" in mathematics. Ten authors were cited who have used "flexpoint," "flex-point," "flex point," or "flex" to designate points at which the slope of a curve attains local extrema, and reasons were given why of these the terminology "flexpoint" would seem the preferable form.

19. *On integrability of vector fields*, by M. Z. Nashed, Georgia Institute of Technology.

The purpose of this paper is to point out the connection between a conservative vector field and the symmetry of the (Fréchet) derivative of the operator inducing the field and to discuss an elementary derivation of the integrability condition in terms of bilinear functionals. A simple procedure for numerical integration of such fields using a scalar quadrature is obtained as a by-product. The procedure does not utilize any coordinate representation of the field and is useful in situations where only a discrete representation of the vector field is given.

20. *A Markov limit process involving Fibonacci numbers*, J. D. Neff, Georgia Institute of Technology.

The two stage Markov transition matrix

$$P = \begin{pmatrix} 0 & 1 \\ a & 1-a \end{pmatrix}$$

has fixed probability vector $\alpha = (\alpha_1, \alpha_2)$, where α_i are limiting expected values. Deleting all branch probabilities from the tree diagram, analogous limits are found, using the limit of the ratio of consecutive elements of Fibonacci sequence. If $a = (\sqrt{5}-1)/2$, both sets of limits are the same.

21. *A case of illogic*, by C. G. Phipps, Tennessee Technological University.

In the proofs usually given for establishing the law of cancellation and the uniqueness of the additive inverse, it is found that one of these theorems is needed to establish the other. Hence, unless one of them is taken as an assumption, the reasoning is circular. The same situation exists in the treatment of cancellation and the multiplicative inverse.

22. *On submodules (subspaces) which determine certain linear (continuous) mappings*, by J. C. Pleasant, East Carolina College.

The topological part of the following statement appeared in this MONTHLY as a problem proposed by A. Weinstein (5232, (1965, 923)): Hausdorff spaces (torsion-free modules over a commutative integral domain A) are characterized as spaces (A-modules) Y satisfying (P). For any topological space (A-module) X, two continuous (A-linear) functions from X into Y which agree on a dense subset (an essential submodule) of X are equal.

Simple examples show that dense subsets and essential submodules are not characterized by (P). Those submodules which are characterized by (P) are discussed and a generalization of the algebraic part is given.

23. *Certain differential equations with a family of almost periodic solutions*, by T. G. Proctor, Clemson University.

This paper gives a theorem which guarantees the existence of a family of almost periodic solutions for the vector differential equation $\dot{y}=\epsilon f(t, y)$ when ϵ is sufficiently small, where $f(t, y)=-f(-t, y)$ is analytic and almost periodic with base of frequencies $\{\omega_1, \omega_2, \cdots, \omega_n\}$ and $\left| \sum_1^n k_i\omega_i\right| \geqq K/(\sum_1^n |k_i|)^{n+1}$ for some $K>0$ and all integral vectors $(k_1, \cdots, k_n)\neq 0$.

24. *Numerical experiments in mini-max approximations*, by R. G. Selfridge, University of Florida.

The main problem in mini-max approximations of more than one independent variable is that the algorithms work reliably only when quite close to a final solution. This paper describes some numerical experiments in such approximations.

25. *The bound-one projection property for hyperplanes*, by Robert Silber, Clemson University.

This presentation was a preliminary report of work on the following problems. Say that a Banach space $X\in P_1^1$ if it has the property that, for each normed linear superspace $Z\supset X$ in which X has codimension 1, there exists a projection p of Z onto X having bound 1. The space $X\in P_1$ if there is a projection of bound 1 onto X from every Banach superspace $Z\supset X$; a separable $X\in S_1$ if there is a projection of bound 1 onto X from every separable Banach superspace Z. If $X\in P_1^1$, then must it be true that $X\in P_1$? If X is separable and $X\in S_1$, then does it follow that $X\in P_1$? (Sobczyk proved that the space (c_0) does not have the projection property, but that it is true that $(c_0)\in S_2$.)

26. *The relative-logarithmic transformation*, by H. Slaughter, Jacksonville State College.

Data (6000 dart throws and 5000 weather temperatures) showed that it was possible to disprove the identity transformation $(x=v)$, which is the normal distribution independent variable, and the log-normal transformation $(x=\ln v)$, but it was not possible to disprove (at the 99% confidence level) the relative-logarithmic transformation $(x=\ln [(\ln v-\ln v_0)/(\ln v_m-\ln v)])$, where v is a physical measurement (absolute temperature, dart reading from center of target), and v_0 and v_m are the lower and upper limits. This transformation fits over 50 distributions of real data, and it suggests that space may be apparently relative-logarithmic to the observer.

27. *Initial stresses in a cylindrically aeolotropic material*, by C. B. Smith, University of Florida.

A method of determining initial stresses in a cylindrically aeolotropic material is developed. The applicability of the method to a pole consisting of the entire section of the trunk of a tree lacking a small circular cylinder containing the longitudinal axis is discussed. The initial stresses in the pole are produced by a change in moisture content.

28. *An algorithmic approach to finitely generated abelian groups*, by D. A. Smith, Duke University.

An algorithm is given for computing the structure, as a direct sum of cyclic groups, of a finitely generated abelian group given by generators and relations. The algorithm is used to give constructive proofs of all the standard theorems about these groups: the basis theorem, invariance of rank, canonical and primary bases. Examples of both hand and machine computations are included, and the algorithm is given as an ALGOL procedure.

29. *Graph coloring and combinatorial numbers*, by Andrew Sobczyk Clemson University.

For a discussion of the Ramsey numbers $N(p, q, 2)$, see Ryser's Carus Monograph, pp. 38–43. In this paper, several new combinatorial numbers are introduced, and their values are determined. For example, the three-color number $M(p, q, 2)$ is the smallest integer such that if $n\geqq M(p, q, 2)$, the n-configuration is sure to contain either a one-colored p-tuple, or a two-colored q-tuple. It is shown that: $M(p, 3, 2)=5$, $p\geqq 3$; $M(3, 4, 2)=8<N(3, 4, 2)=9$; $M(3, 5, 2)=14=N(3, 5, 2)$; $M(3, 6, 2)=17\leqq N(3, 6, 2)$. Greenwood and Gleason determined that $N(3, 6, 2)$ is 17, 18, or 19; in this paper it is shown that $N(3, 6, 2)$ is 17 or 18.

30. *Maximal connected topological spaces*, by J. P. Thomas, The University of North Carolina at Charlotte.

The principal result is THEOREM 4: Let (X, T_1) be a topological space where X has at least two elements and T_1 is such that every intersection of open sets is open. Let I be the set of all isolated points of (X, T_1) and let $J=X-I$. If $x\in J$ let Vx be the smallest neighborhood of x (i.e. the intersection of all open neighborhoods of x). Then in order that (X, T_1) be maximal connected it is necessary and sufficient that all of the following three statements be true:

(i) $\bigcup_{x\in J} Vx=X$.
(ii) If $x\neq x'$, x and $x'\in J$, then $Vx\cap Vx'$ has at most one point.
(iii) If $a, b\in(X, T_1)$ then there exists exactly one simple chain from a to b of open sets Vx.

31. *Stress distribution of a rotating epitrochoid of two semi-cusps*, by J. L. Tilley, Mississippi State University.

Formulae are derived for the calculation of the stress components of a rotating plate of isotropic material in the form of an epitrochoid of two semi-cusps about an axis through its center perpendicular to the plane of the plate.

32. *A theorem on semigroups*, by A. D. Wallace, University of Florida.

A new and simpler proof is given of a theorem of Aczel-Wallace, which generalizes a theorem of Hosszu. Let $T\times X\rightarrow X$ be a continuous function whose value at (t, x) is denoted by tx and let T and X be nonvoid Hausdorff spaces. If the condition $t(t'x)=t'(tx)$ holds for all $t, t'\in T$ and all $x\in X$, if $Ta=\{ta \mid t\in T\}=X$ for some $a\in X$, and if each of T, X is compact or discrete, then there is a continuous associative and commutative operation $\circ$ on X such that $t(x\circ y)=(tx)\circ y$ for all $t\in T$ and all $x, y\in X$. Moreover a is a unit for $\circ$.

HENRY SHARP, JR., *Secretary-Treasurer*

MARCH MEETING OF THE SOUTHERN CALIFORNIA SECTION

The forty-sixth meeting of the Southern California Section of the Mathematical Association of America was held at Occidental College, Los Angeles, California, on March 12, 1966. The registered attendance was 148, including 132 members of the Association. Professor T. M. Apostol, Chairman of the Section, presided at the morning and afternoon sessions.

At the business meeting Professor P. A. White, Chairman of the Nominating Committee, reported the election of the following officers to serve for the year beginning July 1, 1966: Chairman, Professor V. C. Harris, California State College at San Diego; Vice-Chairman, Professor F. A. Valentine, University of California, Los Angeles. The Secretary is completing the second year of a three-year term which ends in June of 1967. The following members of the Program Committee for the 1967 meeting were also elected: Professor E. I. Deaton, San Diego State College, Chairman; Professors C. B. Tompkins and H. G. Tucker.

Professor R. C. James reported as Governor for the Section, giving details of some actions taken by the Board of Governors at the January, 1966, meeting. A report of activities of the High School Mathematics Contest Committee was submitted by the chairman of the committee, Dr. J. M. Huffman. Participation on the part of high schools and high school students increased about 40% from 1965 to 1966.

Awards were made to twelve students from colleges in the Section who placed high in the list of winners of the two most recent William Lowell Putnam Examinations; winners included the first-prize team in the 1964 Contest, from California Institute of Technology, and individual winners from Cal Tech, Pomona College, and U.C.L.A.

The following program was presented:

1. *A generalized Bairstow algorithm*, by Thomas Robertson, Occidental College.

An algorithm of Bairstow type is defined which produces a factor $P_2-\alpha P_1-\beta P_0$ for a linear

combination of polynomials $P_n(x)$ which satisfy a three-term recursion. A theorem is established indicating that the convergence of the successive approximations is normally quadratic. Application of the algorithm to root finding and to the eigenvalue problem for arbitrary tridiagonal matrices is discussed.

2. *Solvable and unsolvable problems*, by Yiannis Moschovakis, University of California, Los Angeles, introduced by the secretary.

In this expository talk, Hilbert's tenth problem (to determine whether an arbitrary Diophantine polynomial has integral roots) is discussed at some length, since it is the most outstanding open problem which is conjectured to be unsolvable. The emphasis is on explaining the nature of unsolvable problems, through a discussion of recursive functions and Church's thesis. An outline is given of S. C. Kleene's proof of the Gödel Incompleteness Theorem, a proof which depends on the existence of some unsolvable arithmetical problem.

3. *The notion of random process*, by Julius Bendat, Measurement Analysis Corporation, a lecture presented jointly by the SIAM and the MAA.

This paper presents many aspects of random processes. The treatment includes a discussion of fundamental probability and statistical considerations underlying random data. Details from practical analog and digital computer techniques for measuring basic parameters of stationary and nonstationary data are given. Recent results are noted concerning various advanced data processing procedures and associated error analyses. Material for this paper is taken from a new book by J. S. Bendat and A. G. Piersol, *Measurement and Analysis of Random Data*, Wiley, New York, 1966.

4. *Theory of distributions and generalized functions*, by Bernard Gelbaum, University of California, Irvine, introduced by the secretary.

An exposition of the basic theory of distributions and generalized functions and their applications to the theory of partial differential equations.

5. *A role for the computer in a liberal arts college*, by John Ferling, Claremont Men's College.

Assuming that a liberal arts college wants to prepare its students to take a place of responsibility in society, the author of this paper affirms that this preparation should include some understanding of digital computers. The computer is rapidly becoming an integral part of our civilization and produces changes comparable to those produced by the wheel and the steam engine at their time. The veil of mystery surrounding the "electronic brain," a name which is misleading in itself, can be dispelled for the liberal arts student by a sequence of selected programs displaying the strengths and weaknesses of the computer.

R. B. HERRERA, *Secretary*

APRIL MEETING OF THE KENTUCKY SECTION

The forty-ninth annual meeting of the Kentucky Section of the Mathematical Association of America was held at the University of Kentucky, Lexington, Kentucky on Friday and Saturday, April 29 and 30, 1966 with Professor J. H. Wells presiding.

There were 108 people registered for the meeting.

An invited address was given Friday afternoon by Professor B. E. Rhoades of Indiana University. Professor Rhoades spoke on "Background for the General Curriculum for Mathematics in Colleges (GCMC) Report." Also, at the Friday afternoon session, Brother Emeric, St. Xavier High School, Louisville, discussed the effects of the GCMC Report on high school mathematics. On Saturday Professor Donald Babbitt, of the University of California at Los Angeles and Indiana University, spoke on "The Correspondence principle viewed as a problem in differential calculus."

The following officers were elected for 1966–67: Chairman, M. G. Carman, Murray State University, Murray, Kentucky; Secretary-Treasurer, J. C. Eaves, University of Kentucky; Contest Chairman of the Annual High School Mathematics Contest, Leland Scott, University of Louisville.

The following papers were presented at the Saturday session:

1. *Matrices all of whose powers lie close to the identity*, by R. H. Cox, University of Kentucky.

If A is an $n\times n$ matrix of complex numbers, then regard A as a linear transformation from the set of n-tuples of complex numbers into itself and define $\|A\|=\sup_{|x|\leq 1}|Ax|$. If I denotes the $n\times n$ identity matrix, then we have the following theorem: If α is a number in (0, 1) and A is an $n\times n$ matrix such that, for each positive integer p, $\|A^p-I\|\leq\alpha$; then A is I.

2. *The probability of winning a series*, by John Mack, University of Kentucky.

3. *On a question concerning curvature and arc length*, by H. G. Robertson, University of Kentucky.

Let K_m be a disc of positive radius M, centered at the origin with circular boundary C_m, and Γ be a smooth curve of length l such that: (i) at its initial point Γ is tangent to C_m; (ii) Γ has minimum radius of curvature m. THEOREM. If $l<m\pi$, then Γ does not intersect K_m, and if $l>m\pi$, Γ may intersect K_m.

4. *Strictly solvable groups*, by T. K. Seo, University of Kentucky.

5. *The correspondence principle viewed as a problem in differential calculus*, by Donald Babbitt, University of California at Los Angeles and Indiana University.

This was an expository talk discussing the impossibility of assigning differential operators (quantum observables) to classical mechanical observables in such a way as to preserve the usual bracket formalism. In this talk the author limited himself to preserving brackets between the Hamiltonian and a rather limited set of observables.

W. C. ROYSTER, *Secretary-Treasurer*

APRIL MEETING OF THE MICHIGAN SECTION

The annual meeting of the Michigan Section of the Mathematical Association of America was held on Saturday, April 2, 1966 at Wayne University, Detroit, Michigan, in conjunction with the meeting of the Michigan Academy of Arts and Letters. Professor Paul Zwier of Calvin College presided at the sessions. A total of 101 persons attended the meeting.

At the business meeting, the following officers were elected: Professor J. H. Powell, Western Michigan University, Chairman; Professor Beauregard Stubblefield, Oakland University, Vice-Chairman; Professor L. H. Serier, Central Michigan University, Secretary-Treasurer.

Professor K. J. Folley of Wayne University gave a report of the activities of the Board of Governors. The report of the 9th Annual Michigan Prize Competition was given by Professor J. H. McKay, of Oakland University. A copy of the report is on file with the Secretary.

Professor McKay moved that $1000 of the mathematics prize money receipts be allocated to the administration of the Mathematics Prize Competition under the advisement of the board of directors; the motion was seconded and carried.

Professor McKay was commended for his excellent work in administrating this project.

Professor Powell moved that the annual dues be made $1.00 per year; this was also seconded and carried.

The following papers were presented:

1. *A general curriculum in mathematics for colleges*, by Ralph Boas, Northwestern University.

2. *Computer science in the Undergraduate Curriculum*, by Walter Hoffman, Wayne State University.

3. *Statistics in the Undergraduate Curriculum*, by Leo Katz, Michigan State University.

4. *Applied mathematics in the Undergraduate Curriculum*, by Robert Thrall, The University of Michigan.

LESTER H. SERIER, *Secretary-Treasurer*

A NEW SERIES OF MATHEMATICS FILMS

A new series of mathematics films, entitled MATHEMATICS TODAY, is now available. It consists of films produced by the Individual Lectures Project of the MAA's Committee on Educational Media, the films produced by the earlier MAA Committee on Production of Films, and three television programs on mathematics produced by the "Science and Engineering TV Journal" and WNDT in New York.

All films are 16 mm. sound films, many in color. Each film can be leased for a long period or rented for a short period. For the special convenience of mathematics clubs and similar organizations, a series of any ten or more films may be rented at a special reduced rate for screenings one each week on a fixed day.

The series MATHEMATICS TODAY consists of the following films at present:

Basic Films—little or no mathematical knowledge required

Let us Teach Guessing: A Demonstration with GEORGE POLYA. (Color, 60 minutes)
Mathematical Induction: A Lecture by LEON HENKIN. (Color, 60 minutes)
Mr. Simplex Saves the Aspidistra, with FRANK KOCHER, LEON HENKIN, and JULIUS H. HLAVATY. (Color, 30 minutes)
Theory of Limits (Part I—Limits of Sequences): A Lecture by E. J. MCSHANE. (B&W, 34 minutes)
Topology, with RAOUL BOTT and MARSTON MORSE. (B&W, 30 minutes)

Collegiate Films—for the undergraduate & teacher

Challenge in the Classroom: The Methods of R. L. MOORE. (Color, 55 minutes)
Challenging Conjectures: A Lecture by R. H. BING. (B&W, 40 minutes)
Differential Topology: Three Lectures by JOHN MILNOR. (B&W, 60 minutes each)
Fixed Points: A Lecture by SOLOMON LEFSCHETZ. (Color, 60 minutes)
Göttingen and New York—Reflections on a Life in Mathematics—RICHARD COURANT. (Color, 43 minutes)
The Kakeya Problem: A Lecture by A. S. BESICOVITCH. (Color, 60 minutes)
Pits, Peaks, and Passes (Part 1). A Lecture on Critical Point Theory by MARSTON MORSE. (Color, 49 minutes)
Predicting at Random: A Lecture by DAVID BLACKWELL. (Color, 42 minutes)
The Search for Solid Ground: A Panel Discussion with MARK KAC, JOHN KEMENY, HARTLEY ROGERS and RAYMOND SMULLYAN. (B&W, 60 minutes)
Theory of Limits (Parts II & III—Limits of Functions and Limit Processes and The Cauchy Criterion for Convergence): A Lecture by E. J. MCSHANE. (B&W, 38 minutes)
What is an Integral? A Lecture by EDWIN HEWITT. (B&W, 60 minutes)
What is Mathematics and how do we Teach it? A Panel Discussion with LIPMAN BERS, SAMUEL EILENBERG, ANDREW GLEASON, HENRY POLLAK, and LEO ZIPPIN. (B&W, 45 minutes)

Advanced Films—for the mathematician

Applications of Group Theory in Particle Physics: A Lecture by FREEMAN DYSON. (B&W 60 minutes)
Can you Hear the Shape of a Drum? A Lecture by MARK KAC. (Color 47 minutes)
Can you Hear the Shape of a Drum? (Complete Version). A Lecture by MARK KAC. (Color, 65 minutes)

The Classical Groups as a Source of Algebraic Problems: A Lecture by CHARLES CURTIS. (B&W, 60 minutes)

Pits, Peaks, and Passes (Part II). A Lecture on Critical Point Theory by MARSTON MORSE. (Color, 26 minutes)

Additional films are expected to be available in the fall of 1966. All films listed above can be obtained for purchase, lease, or rental from MODERN LEARNING AIDS, 1212 Avenue of the Americas, New York, N. Y. 10036.

ANNOUNCEMENT OF L. R. FORD AWARDS

At its meeting on January 27, 1965, in Denver, Colorado, the Board of Governors authorized a number of awards, to be named after Lester R. Ford, Sr., to authors of expository articles published in the MONTHLY and the MATHEMATICS MAGAZINE. A maximum of six awards will be made annually; each award is in the amount of $100. The articles are to be selected by a subcommittee of the Committee on Publications appointed for this purpose.

The 1966 recipients of these awards, selected by a committee consisting of R. P. Boas, Chairman; C. W. Curtis, and R. P. Dilworth, were announced by President Wilder at the Business Meeting of the Association on August 30, 1966, at Rutgers—The State University. The recipients of the Ford Awards for articles published in 1965 were the following:

C. B. Allendoerfer, Generalizations of Theorems About Triangles, MATH. MAG., 38(1965) 253–259.

P. D. Lax, Numerical Solution of Partial Differential Equations, MONTHLY, 72(1965), Part II (Slaught Paper No. 10) 74–84.

Marvin Marcus and Henryk Minc, Permanents, MONTHLY, 72(1965) 577–591.

HENRY L. ALDER, *Secretary*

THE 1966 WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

The twenty-seventh annual William Lowell Putnam Mathematical Competition will be held on Saturday, November 19, 1966. This competition, made possible by the trustees of the William Lowell Putnam Intercollegiate Memorial Fund left by Mrs. Putnam in memory of her husband, is under the sponsorship of the Mathematical Association of America. Colleges and universities in the United States and Canada are eligible to register undergraduates in the competition.

Application blanks will be mailed about October 1 to the mathematics department chairmen of the schools on the regular mailing list. If an application blank is not received by October 15, one may be secured by writing the director, James H. McKay, c/o Department of Mathematics, University of California, Berkeley. Your application should be filed with the director not later than October 24, 1966. Further details are provided in the announcement which is mailed with the registration forms.

Reports of the previous competitions and the examination questions may be found in the MONTHLY for May 1938, 1939, 1940, 1941, 1942; October 1946; August–September 1947; December 1948; August–September 1949, 1950, 1951; October 1952, 1953, 1954, 1955; December 1956; August–September (announcement of winners) and November (questions and solutions) 1957; August–September 1958, 1959; January (questions and solutions for the eighteenth, nineteenth, and twentieth competitions) 1961; August–September 1961; October 1962; August–September 1963; June–July 1964; August–September 1965; and this issue pages 726-732.

CALENDAR OF FUTURE MEETINGS

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

Forty-eighth Summer Meeting, University of Toronto, Toronto, Ontario, Canada, August 28–30, 1967.

ALLEGHENY MOUNTAIN, West Virginia University, Morgantown, West Virginia, May 6, 1967.

ILLINOIS, University of Illinois, Urbana, May 14–15, 1967.

INDIANA, Purdue University, Lafayette, November 5, 1966.

IOWA, Drake University, Des Moines, April 21, 1967.

KANSAS, Fort Hays State College, Hays, April 22, 1967.

KENTUCKY, Murray State University, Murray, Spring 1967.

LOUISIANA-MISSISSIPPI, Jung Hotel, New Orleans, Louisiana, March 4–5, 1967.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN

MINNESOTA, University of Manitoba, Winnipeg, Manitoba, Canada, October 29, 1966.

MISSOURI, Northeast Missouri State Teachers College, Kirksville, April 29, 1967.

NEBRASKA, University of South Dakota, Vermillion, May 6, 1967.

NEW JERSEY, Rutgers, The State University, New Brunswick, November 12, 1966.

NORTHEASTERN, Trinity College, Hartford, Connecticut, November 26, 1966.

NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.

OHIO

OKLAHOMA-ARKANSAS, Northeastern State College, Tahlequah, Oklahoma, March-April, 1967.

PACIFIC NORTHWEST, University of Montana, Missoula, June 16–17, 1967.

PHILADELPHIA, Villanova University, Villanova, November 19, 1966.

ROCKY MOUNTAIN

SOUTHEASTERN, Florida Presbyterian College, St. Petersburg, Florida, March 31–April 1, 1967.

SOUTHERN CALIFORNIA, San Diego State College, San Diego, March 18, 1967.

SOUTHWESTERN

TEXAS, Austin College, Sherman, April 14–15, 1967.

UPPER NEW YORK STATE, State University College, Plattsburgh, May 20, 1967.

WISCONSIN, St. Norbert College, DePere, May 6, 1967.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, Houston, Texas, January 24–27, 1967.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Michigan State University, June 19–23, 1967.

ASSOCIATION FOR COMPUTING MACHINERY, Sheraton-Park, Washington, D. C., August 29–31, 1967.

ASSOCIATION FOR SYMBOLIC LOGIC, Houston, Texas, January 23–24, 1967.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Houston, Texas, January 28, 1967.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Jack Tar Hotel, Durham, North Carolina, October 17–19, 1966.

ALLYN AND BACON, INC.

Dept. E, 150 Tremont St., Boston, Mass. 02111

paperbound series—

THE ALLYN AND BACON SERIES, TOPICS IN CONTEMPORARY MATHEMATICS:

ALGEBRA THROUGH PROBLEM SOLVING by A. P. Hillman, University of New Mexico and G. L. Alexanderson, University of Santa Clara. 1966. Paperbound. $5\frac{5}{8} \times 8\frac{3}{8}$. 129 pp. $2.95.

SETS WITH APPLICATIONS by Peter W. Zehna, United States Naval Postgraduate School. 1966. Paperbound. 6 x 9. 153 pp. $2.95.

INFORMAL LOGIC by John W. Kenelly, Clemson University. Fall, 1966. Paperbound. $5\frac{5}{8} \times 8\frac{3}{8}$. Est. 128 pp. Tent. $2.95.

a proven success—

CALCULUS WITH ANALYTICAL GEOMETRY, THIRD EDITION by Richard E. Johnson, University of New Hampshire and Fred L. Kiokemeister, Mount Holyoke College. 1964. $6\frac{1}{2} \times 9$. 798 pp. $11.95.

[redacted]

recently published—

ADVANCED CALCULUS: AN INTRODUCTION TO APPLIED MATHEMATICS, Volumes One and Two by Arthur E. Danese, State University of New York at Buffalo.

■ A recent publication in two volumes, the text presents mathematical topics beyond the elementary calculus. Provides the necessary mathematics background for mathematicians, engineers and scientists. The material is treated in a complete manner within a modern framework, yet the author has also kept in mind limitations of students at this level.

■ Contents: VOLUME ONE: Summary of Elementary Mathematics. I. Vector Analysis. II. Complex Analysis. III. Fourier Analysis. IV. Special Functions. VOLUME TWO: Summary of Elementary Mathematics. V. The Calculus of Variations. VI. Operational Analysis. VII. Partial Differential Equations and Boundary Value Problems.

■ 1965. 6 x 9. Volume One: 558 pp. $10.75. Volume Two: 373 pp. $8.25.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 8

CONTENTS

OCTOBER 1966

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453. Manuscripts of papers should be submitted to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE TO AUTHORS

The MONTHLY welcomes papers presenting valid mathematics, of rather general interest, at a level intelligible to persons with two years of full-time graduate study. Some novelty of content, viewpoint, or arrangement is essential. Expository articles are particularly desired. State the context and the principal aim of the paper early. Address yourself quite explicitly to the reader described above, communicating your ideas to him clearly and attractively.

The title should be brief and meaningful. Since the title will be quoted and reproduced by laymen, it should contain no symbols unfamiliar to laymen.

Articles should be typewritten, double-spaced, on $8\frac{1}{2}\times 11''$ paper of very good quality. Submit the original (and a duplicate if convenient) keeping a complete copy for yourself. To avoid loss and delay notify us of any change of address.

The typescript should be prepared with extreme care. Misprints are highly obnoxious; so are dangling participles. Put name and address between title and text. Put references at the end with bracketed citations in the text. Avoid footnotes; instead, use clearly designated remarks in the text. Put acknowledgments at the very end, just before the bibliography. Be generous with spacing and displays. *Keep notation simple.* For a matrix the notation $[a_{jk}]$ is recommended, with det $[a_{jk}]$ or $|a_{jk}|$ for the determinant. On doubtful questions regarding format or notation, observe practices in current issues of the MONTHLY, or consult the applicable sections of the "Author's Manual" of the American Mathematical Society.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

TOPOLOGICAL ASPECTS OF SYLVESTER'S THEOREM ON THE INERTIA OF HERMITIAN MATRICES

HANS SCHNEIDER, University of Wisconsin, Madison

1. A set of $n\times n$ matrices with complex elements has a natural topology associated with it. One may therefore look for a topological interpretation of some results in the theory of matrices. We shall show that Sylvester's classical theorem on the inertia (signature) of Hermitian matrices concerns the connected components of the space of all Hermitian matrices of fixed rank r.

Most of the arguments used in the proof of our theorem are elementary and familiar. Yet our result does not appear in the literature. The reason may well be that matrix theorists tend to use "continuity properties" as they arise, without formalizing them, while topologists do not usually study equivalence relations on matrices. This note is offered as an illustration that even on a fairly elementary level, something is gained by looking for inter-connections between different mathematical fields.

2. Let n be a positive integer and let $\omega=(\pi, \nu, \delta)$ be an ordered triple of nonnegative integers with $\pi+\nu+\delta=n$. Let E_ω be the diagonal matrix $E_\omega=\operatorname{diag}(1, \cdots, 1, -1, \cdots, -1, 0, \cdots, 0)$ with π diagonal elements 1, ν diagonal elements -1 and δ elements 0. Sylvester's theorem ([7] p. 100, [3] p. 83) on the inertia of Hermitian matrices asserts that for each $n\times n$ Hermitian matrix H there exists just one matrix E_ω for which there exists a nonsingular matrix X such that $X^*HX=E_\omega$.

But can we pick out the triple ω that occurs in Sylvester's theorem, without use of that theorem and directly in terms of the matrix H? One possibility, which we mention merely because of its intrinsic interest, is to proceed geometrically. Let V be the space of all positive n-tuples and associate with H the quadratic form Δ: $(x, x)=x^*Hx$. If it should happen that for some $y\in V$, $(y, y)>0$ then also $(x, x)>0$ for $x=\alpha y$ if $\alpha\neq 0$. Thus we have found a subspace W of V of which $(x, x)>0$ for all $x\neq 0$. In other words, Δ is positive definite on W. Now suppose that π is the dimension of a subspace W of largest dimension on which Δ is positive and similarly suppose ν is the dimension of a subspace W' of largest dimension on which Δ is negative definite (i.e., $(x, x)<0$, if $0\neq x\in W'$). If $\delta=n-\pi-\nu$, then it may be proved that $\omega=(\pi, \nu, \delta)$ is the ω of Sylvester's theorem, (see [1] pp. 148–150).

3. We shall use an entirely different approach. Since H is Hermitian all its eigenvalues are real. We shall define π, ν, δ in terms of the eigenvalues of H.

DEFINITION 1. *Let $\pi(H)=\pi$ be the number of positive eigenvalues of H, $\nu(H)=\nu$ the number of negative eigenvalues of H, and $\delta(H)=\delta$, the number of zero eigenvalues of H. Then the ordered triple $\omega=(\pi, \nu, \delta)$ will be called the inertia of H. We shall write $\omega=\operatorname{In} H$.*

DEFINITION 2. *Two Hermitian matrices H, K are inertially equivalent if* $\text{In } H = \text{In } K$. *We shall write* $H \overset{i}{\sim} K$.

The next definition is standard.

DEFINITION 3. (e.g. [6] p. 99, p. 84). *Two Hermitian matrices H and K are conjunctive (conjunctively equivalent) if there exists a nonsingular X such that* $X^*HX = K$. *We shall write* $H \overset{c}{\sim} K$.

It is evident that $\overset{i}{\sim}$ and $\overset{c}{\sim}$ are equivalence relations on any set of Hermitian matrices.

4. Our next two equivalence relations are of a different kind, since they may be defined on any topological space.

If E is a topological space, the space is called connected if the empty set and E are the only subsets of E which are both open and closed (see [5] p. 117). A subset U of E is connected if and only if it is a connected space in the topology induced on U by E. Thus U is connected if and only if, for any set $F \subseteq E$ which is both open and closed, either $U \subseteq F$ or $U \subseteq E \backslash F$, the complement of F in E. This motivates

DEFINITION 4. *Let E be a topological space. We call $x, y \in E$ connectable in E if for every open and closed set U in E both $x, y \in U$ or both $x, y \notin U$. We shall write* $x \overset{u}{\sim} y$.

An *arc* in a topological space E is a continuous image of the interval (0, 1) on the real line in the space E ([5] p. 139).

DEFINITION 5. *Let E be a topological space. We call $x, y \in E$ arc connectable in E if there exists an arc in E joining x, y, i.e., if there exists a continuous function f from the unit interval* (0, 1) *on the real line into E with $f(0) = x$ and $f(1) = y$. We shall write* $x \overset{a}{\sim} y$.

The following lemma is a restatement of a well-known result (see [5] p. 141).

LEMMA 1. *If E is a topological space, then $x \overset{a}{\sim} y$ implies that $x \overset{u}{\sim} y$.*

Proof. Let U be any open and closed set containing x and let f be a continuous function of (0, 1) into E with $f(0) = x$ and $f(1) = y$. Then $f^{-1}(U)$ is an open and closed subset of (0, 1) which contains 0, and the only such set is (0, 1) itself. Thus $y = f(1) \in U$ and so $x \overset{u}{\sim} y$.

5. Let S be any set of $n \times n$ matrices. We can norm S in many ways. For example we can put $\|A\| = \max_{i,j} |a_{ij}|$ for $A \in S$. To turn S into a topological space we choose as the open sets arbitrary unions of finite intersections of all cubes $N(A, \epsilon) = \{B \in S: \max_{i,j} |b_{ij} - a_{ij}| = \|B - A\| < \epsilon\}$, with $A \in S$ and $\epsilon > 0$. Thus a subset T of S is open if and only if for $A \in T$ we can find on $\epsilon = 0$ such that $N(A, \epsilon) \subseteq T$. Observe that S need not be a linear space, nor will our topological space S necessarily be complete.

In this section, we shall consider the space N of all nonsingular complex matrices normed as above.

LEMMA 2. *For all $A, B \in N$, A is arc connectable to B in N.*

Proof. It is enough to prove that for all $A \in N$, $A \overset{a}{\sim} I$, the identity matrix. Choose σ so that $e^{i\sigma}A$ has no negative eigenvalue, and set $f(t) = e^{i\sigma t}A$, $0 \leqq t \leqq 1$. If A belongs to N, so does $e^{i\sigma t}A$ and clearly f is continuous. Thus $A \overset{a}{\sim} C$. Next set $g(t) = (1-t)C + tI$, $0 \leqq t \leqq 1$. Evidently g is again continuous, and since the eigenvalues of $g(t)$ are of the form $\gamma(t) = (1-t)\gamma + t$ where γ is an eigenvalue of C and here γ is not negative, it follows that $\gamma(t) \neq 0$ and so $g(t)$ is nonsingular. Thus $C \overset{a}{\sim} I$. Hence $A \overset{a}{\sim} I$, and this completes the proof.

6. We now require a lemma of a different type. Usually it is expressed by asserting that the eigenvalues of a matrix are continuous functions of the elements of the matrix. We shall state the result precisely:

LEMMA 3. *Let A be a matrix with distinct eigenvalues $\alpha_1, \cdots, \alpha_s$ of multiplicities $m_1, \cdots, m_s$ respectively. Let $\epsilon > 0$ and let $\Gamma(\alpha_i, \epsilon)$ be the circle with center α_i and radius ϵ. Then there is a positive σ, such that every matrix B, for which $\|B-A\| < \sigma$, has exactly m_i eigenvalues in the circle $\Gamma(\alpha_i, \epsilon)$.*

Proof. Let $p(t) = \lambda^n + p_{n-1}\lambda^{n-1} + \cdots + p_0$, and $q(t) = \lambda^{n-1} + q_{n-1}\lambda^{n-1} + \cdots + q_0$. We use a theorem on the zeros of a polynomial (see [4], p. 3). If the zeros of $p(\lambda)$ are α_i with multiplicity m_i, $i = 1, \cdots, s$, $\alpha_i \neq \alpha_j$, if $i \neq j$, and if $(q_j - p_j) < \eta$, $j = 0, 1, \cdots, n-1$, where η is sufficiently small, then m_i zeros of $q(t)$ lie in the circle $\Gamma(\alpha_i, \epsilon)$. Now the eigenvalues of A and B are simply the zeros of the characteristic polynomials det $(\lambda I - A)$ whose coefficients are sums of products of elements of A, and similarly for B. Since addition and multiplication of complex numbers is continuous, we deduce that for sufficiently small $\sigma > 0$, $\|B-A\| < \sigma$, implies that $|q_i - p_i| < \eta$, $j = 0, \cdots, n-1$, and the result follows.

Our proof of Lemma 3 is not really much of a proof, since it refers the result for the spectra of a matrices back to the corresponding theorem for the zeros of polynomials ("continuity of zeros of polynomials"). This latter result is deeper than any other theorem we have used in this note, and we shall not attempt to prove it here. We may note that in the application of Lemma 3, the matrices A and B are both Hermitian. For normal, and therefore for Hermitian matrices, a more precise result is given in [2]:

LEMMA 4. *If A and B are normal matrices with eigenvalues $\alpha_1, \cdots, \alpha_n$ and $\beta_1, \cdots, \beta_n$ respectively, then there exists a suitable numbering of the eigenvalues such that*

$$\sum_i |\alpha_i - \beta_i|^2 \leqq \sum_{i,j} |a_{ij} - b_{ij}|^2.$$

This result rests on the famous theorem of Birkhoff that the permutation matrices are the vertices of the convex polyhedron of doubly stochastic matrices. For a proof of Birkhoff's theorem see [3] p. 97, or [2], and for a proof of

Lemma 4 see [2]. Of course, Lemma 4 implies Lemma 3 since $\sum_{i,j} |a_{ij}-b_{ij}|^2 \leqq n^2\|A-B\|^2$.

7. From now on our space will be the space H_r^n of all $n\times n$ Hermitian matrices of *fixed rank r*. We shall first examine a trivial situation. The space of all Hermitian 1×1 matrices is just the real line R and hence H_1^1 is the real line with the origin removed. The connectivity properties of R and H_1^1 are quite different. R is connected and H_1^1 is not. Similarly the connectivity properties of H_r^n will be quite different from these of the space of all $n\times n$ Hermitian matrices. The reason for focusing on H_r^n is that this space yields an interesting theorem.

Notation. Let $H\in H_r^n$. Then the set of all $K\in H_r^n$ such that $H\overset{i}{\sim}K$ will be denoted by $\mathbf{I}(H)$, and the eigenvalue class $\mathbf{I}(H)$ will be called an inertial component of H_r^n. Similarly we define $\mathbf{C}(H)$, $\mathbf{U}(H)$, $\mathbf{A}(H)$ to be the equivalence classes of H for $\overset{c}{\sim}$, $\overset{u}{\sim}$, $\overset{a}{\sim}$, respectively, and we call $\mathbf{C}(H)$ a conjunctive component, $\mathbf{U}(H)$ a connected component and $\mathbf{A}(H)$ an arc component of H_r^n.

THEOREM. *Let H_r^n be the topological space of all $n\times n$ Hermitian matrices of rank r. Then the four equivalence relations $\overset{i}{\sim}$, $\overset{c}{\sim}$, $\overset{a}{\sim}$, $\overset{u}{\sim}$, coincide on H_r^n. Equivalently, for any $H\in H_r^n$, $\mathbf{I}(H)=\mathbf{C}(H)=\mathbf{A}(H)=\mathbf{U}(H)$.*

Proof. We shall prove that $H\overset{i}{\sim}K$ implies $H\overset{c}{\sim}K$, $H\overset{c}{\sim}K$ implies $H\overset{a}{\sim}K$, $H\overset{a}{\sim}K$ implies $H\overset{u}{\sim}K$, and $H\overset{u}{\sim}K$ implies $H\overset{i}{\sim}K$.

(a) $H\overset{i}{\sim}K$ implies $H\overset{c}{\sim}K$: Suppose $\operatorname{In} H=\operatorname{In} K=\omega=(\pi,\ \nu,\ \delta)$ say. It is enough to prove that $H\overset{c}{\sim}E_\omega$, where E_ω is defined in Section 1. (Note that $\pi+\nu=r$ and that $E_\omega\in H_r^n$.) Since H is Hermitian there exists a unitary Y such that $Y^*HY=\operatorname{diag}(\alpha_1,\cdots,\alpha_n)$. By definition of H, and since the α_i are the eigenvalues of H, it follows that π of the α_i are positive, ν are negative and the remaining δ are zero. Replacing Y, if necessary, by the unitary matrix YP, where P is a permutation matrix, we may suppose that $Y^*HY=\operatorname{diag}(\alpha_1,\cdots,\alpha_n)$ where $\alpha_i>0$, $i=1,\cdots,\pi$, $\alpha_{i+1}<0$, $i=\pi+1,\cdots,\pi+\nu$, and $\alpha_i=0$, $i=\pi+\nu+1,\cdots,n$.

Thus if $D=\operatorname{diag}(\sqrt{\alpha_1},\cdots,\sqrt{\alpha_\pi},\sqrt{-\alpha_{\pi+1}},\cdots,\sqrt{-\alpha_{\pi+\nu}},1,\cdots,1)$ and $X=YD^{-1}$, then $X^*HX=E_\omega$.

(b) $H\overset{c}{\sim}K$ implies $H\overset{a}{\sim}K$: Let $X^*HX=K$, where X is nonsingular. Since by Lemma 2 any two nonsingular matrices are connected in the space N of nonsingular complex matrices, there is a continuous function $t\rightarrow X(t)$ of $(0,1)$ into N such that $X(0)=I$ and $X(1)=X$. Further, transposition and matrix multiplication are continuous operations. Hence the function $f(t)=X^*(t)HX(t)$ is continuous. But rank $X^*(t)HX(t)=r$, since $X(t)$ is nonsingular, whence $X(t)HX(t)\in H_r^n$. Further $f(0)=H$ and $f(1)=K$, and so $H\overset{a}{\sim}K$.

(c) $H\overset{c}{\sim}K$ implies $H\overset{u}{\sim}K$. This is just Lemma 1.

(d) $H\overset{u}{\sim}K$ implies $H\overset{i}{\sim}K$. We shall prove the equivalent result that $\mathbf{I}(H)$ contains $\mathbf{U}(H)$, for $H\in H_r^n$: Let $K\in\mathbf{I}(H)$, and let $\operatorname{In} K=\omega=(\pi,\ \nu,\ \delta)$, and suppose the eigenvalues α_i, $i=1,\cdots,\pi$, of K are positive, the eigenvalues α_i,

$i=\pi+1, \cdots, \pi+\nu$ are negative and the eigenvalues α_i, $i=\pi+\nu+1, \cdots, n$ are zero.

Let $0<\epsilon<\min \{|\alpha_i|:\alpha_i\neq 0\}$. By Lemma 3 there exists a neighborhood $N(K, \sigma)$ of K in H_r^n (thus each $L\in N(K, \sigma)$ is, by assumption, Hermitian of rank r) so that the spectrum of each $L\in N(K, \sigma)$ is contained in the union of the n circles $\Gamma(\alpha_i, \epsilon)$. It follows that each L in this neighborhood of K has at least as many positive (negative) eigenvalues as K has positive (negative) eigenvalues, i.e. $\pi(L)\geqq\pi$, and $\nu(L)\geqq\nu$. But as $L\in H_r^n$, $\pi(L)+\nu(L)=r=\pi+\nu$, whence $\pi(L)=\pi$ and $\nu(L)=\nu$. Thus $\text{In } L=\text{In } K$. It follows for each $K\in\mathbf{I}(H)$, there exists an $N(K, \sigma)\subseteq\mathbf{I}(H)$, and so $\mathbf{I}(H)$ is open. Now $H_r^n\backslash\mathbf{I}(H)=\bigcup\{\mathbf{I}(M): M\in H_r^n, M\notin\mathbf{I}(H)\}$ and a union of open sets is open. Hence $\mathbf{I}(H)$ is also closed. By the definition of $\mathbf{U}(H)$, we see that $\mathbf{U}(H)$ is contained in every open and closed set containing H, whence $\mathbf{U}(H)\subseteq\mathbf{I}(H)$. This completes the proof of the theorem.

COROLLARY. *The topological space H_r^n has precisely $r+1$ distinct inertial components (or conjunctive components, or connected components, or arc components).*

Proof. Obviously, each $\omega=(\pi, \nu, \delta)$ with $\pi+\nu=r$ corresponds to one inertial component of H_r^n, and there are just $r+1$ such ω. By the theorem, each inertial component is a conjunctive component (and a connected component and an arc component).

8. It should be noted that in the proof of our theorem (a) is just a standard proof that there exists an ω such that $H\overset{c}{\sim}E_\omega$. The direct proof that this ω is unique is simple (see [7] pp. 92, 100), but we do not need this proof. For distinct ω, the corresponding E_ω obviously lie in distinct inertial components and the uniqueness now follows from the equality of $\overset{c}{\sim}$ and $\overset{i}{\sim}$ on H_r^n.

The concept of inertia may be extended to matrices that are not Hermitian. For some results in this direction see [6].

I wish to thank Marvin Marcus, Judith Molinar, and Robert Thompson for comments which have helped to improve this paper.

References

1. E. Artin, Geometric Algebra, Interscience, New York, 1957.

2. A. J. Hoffman and H. Wielandt, The variation of the spectrum of a normal matrix, Duke M. J., 20 (1953) 37–40.

3. M. Marcus and H. Minc, A survey of matrix theory and matrix inequalities, Allyn and Bacon, Boston, 1964.

4. M. Marden, The geometry of the zeros of a polynomial in a complex variable, Math. Surveys III, American Math. Soc., 1949.

5. B. Mendelson, Introduction to Topology, Allyn and Bacon, Boston, 1962.

6. A. Ostrowski and H. Schneider, Some theorems on the inertia of matrices, J. Math. Analysis Appl., 4(1962) 72–84.

7. S. Perlis, Theory of matrices, Addison-Wesley, Reading, Mass., 1952.

SCRAMBLED SERIES

PAUL JOHNSON AND RAY REDHEFFER, University of California, Los Angeles

1. Introduction. Let a_k be defined for $k=1, 2, 3, \cdots$. A *rearrangement* of $\sum a_k$ is a series $\sum b_k$ where each a_k occurs just once among the b_k's, and *vice versa*. If $a_k \geqq 0$ and $\sum a_k < \infty$ then

$$\sum_1^n b_k \leqq \sum_1^m a_k \leqq \sum_1^\infty a_k$$

as soon as m is so large that every b_k in the first sum occurs among the terms of the second. Hence $\sum b_k$ converges, and $\sum b_k \leqq \sum a_k$. By logical symmetry $\sum a_k \leqq \sum b_k$ and equality holds. Furthermore, when either series diverges, so does the other.

If the a_k change sign but $\sum |a_k| < \infty$ then the identities

$$a_k = (a_k + |a_k|) - |a_k|, \qquad b_k = (b_k + |b_k|) - |b_k|$$

together with the linearity of "$\sum$" enable us to deduce $\sum a_k = \sum b_k$ from the foregoing result. And if $a_k = u_k + iv_k$ is complex, with $\sum |a_k| < \infty$, the fact that $|u_k| \leqq |a_k|$ gives $\sum |u_k| < \infty$. Similarly $\sum |v_k| < \infty$ and hence, $\sum a_k = \sum b_k$ again.

These observations are, of course, familiar from the sophomore curriculum. But suppose we make a more drastic rearrangement, in which infinitely many a_k's are grouped together. For instance, let $b_1 = \sum a_k$ over the prime subscripts k, let $b_2 = \sum a_k$ over $k=4, 6, 8, \cdots$, let $b_3 = \sum a_k$ over the multiples of 3 that have not yet been used, and so on (as in Eratosthenes' sieve [1]). It is as clear as anything can be that $\sum b_k = \sum a_k$, if $a_k \geqq 0$ and yet this trivial fact escapes the foregoing analysis. The series $\sum a_k$ has not just been rearranged; it has been scrambled.

Our interest in the subject arose in 1962, when we gave a course of advanced analysis for high-school teachers. To emphasize the richness and variety of mathematical thought a brisk pace was considered better than a plodding one; e.g., the absolute convergence of single and double series, the Weierstrass approach to function theory via power series, and the development of n-dimensional measure were all covered in just a few periods, by means of the methods described here. Later we noted with pleasure that similar methods are used in Knopp [1]. But there the subject is scattered over two hundred pages; and hence, a major purpose of this expository article is to make Knopp's ideas more readily available. In our opinion these ideas should be part of every course on advanced calculus.

2. A miniature theorem. For precision of statement, we give a formal definition:

DEFINITION 1. *Let the set of integers* $\{1, 2, 3, \cdots\}$ *be partitioned into disjoint nonempty subsets* $S_1, S_2, S_3, \cdots$, *the elements of each* S_k *being arranged in a defi-*

nite linear order. Using this order, form the sums

$$b_1 = \sum_{i\in S_1} a_i, \qquad b_2 = \sum_{i\in S_2} a_i, \qquad b_3 = \sum_{i\in S_3} a_i,$$

and so on. Then $\sum b_k$ is a scrambled form of $\sum a_k$.

The term "partition" implies that $S_1\cup S_2\cup S_3\cup \cdots = \{1, 2, 3, \cdots\}$. An ordinary rearrangement is obtained when each S_k has just one element.

THEOREM 1. *Let $a_k \geqq 0$, or let $\sum|a_k| < \infty$. Then $\sum b_k = \sum a_k$ for every scrambled form, $\sum b_k$.*

The reason for distinguishing the case $a_k\geqq 0$ is to allow the possibility that $\sum a_k = \sum b_k = \infty$, where $\sum b_k = \infty$ means either that this series diverges, or that some series defining a b_k diverges.

For proof consider first the case in which $a_k \geqq 0$ and $\sum a_k < \infty$. Then $\sum \tilde{a}_k < \infty$, where $\tilde{a}_k$ is a_k or 0. Since the series defining b_k is a rearrangement of some such series $\sum \tilde{a}_k$, the remarks of the introduction show that $\sum b_k$ is well determined, independently of the order that was imposed on the elements of S_k.

We now show that $\sum b_k \leqq \sum a_k$. Given $\epsilon > 0$ we can make

$$b_1 < \textstyle\sum_1 a_k + \frac{1}{2}\epsilon, \qquad b_2 < \sum_2 a_k + \frac{1}{4}\epsilon, \qquad b_3 < \sum_3 a_k + \frac{1}{8}\epsilon,$$

and so on, where $\sum_1$ is over a finite subset of S_1, $\sum_2$ is over a finite subset of S_2, and so on. By addition

$$b_1 + b_2 + \cdots + b_m \leqq \textstyle\sum' a_k + \epsilon \leqq \sum a_k + \epsilon,$$

where $\sum'$ is over a finite subset of the a_k. This gives $\sum b_k \leqq \sum a_k$.

On the other hand it is easily seen that

$$\sum_1^n a_k \leqq \sum_1^m b_k \tag{1}$$

as soon as m is large enough and hence $\sum a_k \leqq \sum b_k$. Therefore equality holds.

If $a_k \geqq 0$ and $\sum a_k = \infty$, then (1) gives $\sum b_k = \infty$. Extension to the case in which a_k changes sign or is complex follows as in the introduction.

3. Double series. If $\sum b_k$ and $\sum c_k$ are both scrambled forms of $\sum a_k$ the conclusion of Theorem 1 implies that

$$\sum b_k = \sum c_k. \tag{2}$$

As an illustration we have:

COROLLARY 1. *If $a_{ij} \geqq 0$ then $\sum_{i=1}^\infty \sum_{j=1}^\infty a_{ij} = \sum_{j=1}^\infty \sum_{i=1}^\infty a_{ij}$.*

The assertion allows the case "$\infty = \infty$." For proof let a_i be any enumeration of the a_{ij}, such as the corner clipping enumeration

$$a_1 = a_{11},\ a_2 = a_{12},\ a_3 = a_{21},\ a_4 = a_{13},\ a_5 = a_{22},\ a_6 = a_{31},$$

and so on. If we form the series

$$b_k = \sum_{j=1}^{\infty} a_{kj}, \qquad c_k = \sum_{i=1}^{\infty} a_{ik}$$

then $\sum b_k$ and $\sum c_k$ are both scrambled forms of $\sum a_i$ and the conclusion follows, therefore, from (2).

The essence of the matter is that a partition of the (i, j) lattice induces a corresponding partition on the index-set $\{k\}$ for a_k. We can therefore give a vastly stronger theorem, as follows. Let the plane lattice (i,j) for $i,j=1, 2, 3, 4, \cdots$ be divided in any way whatever into sets $S_1, S_2, S_3, \cdots$, such that each (i, j) is in just one S_k. The sets need not be finite, but in each S_k, the elements are enumerated in some definite order. A *method of summation* of $\sum a_{ij}$ is given by $\sum b_k$, where

$$b_k = \sum a_{ij}, \qquad (i, j) \in S_k,$$

the latter sum being formed in the prescribed order. For example, one might have

$$b_1 = \sum_{i=1}^{\infty} a_{i1}, \qquad b_2 = \sum_{=2}^{\infty} a_{2j}, \qquad b_3 = \sum_{i=3}^{\infty} a_{ii}, \qquad b_4 = a_{56} + a_{78},$$

and so on.

We say that $\sum a_{ij}$ is *absolutely convergent* if $\sum |a_i| < \infty$ for some enumeration a_i of a_{ij}. By the ordinary rearrangement theorem this happens for all enumerations if for any, and the proof of Corollary 1 shows that the condition is equivalent to either of

$$\text{(3a)} \qquad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{ij}| < \infty, \qquad \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |a_{ij}| < \infty.$$

Thus, the notion of absolute convergence is well defined.

COROLLARY 2. *If $\sum a_{ij}$ is absolutely convergent, then all methods of summation give the same value.*

The proof is exactly like the proof of Corollary 1. It is very important that the plane regions S_k need not be bounded—the familiar process of exhausting the (i, j) plane by *finite* regions does not even yield Corollary 1.

An evident consequence of (3a) and Corollary 2 is:

COROLLARY 3. *If $\sum |a_i| < \infty$ and $\sum |b_i| < \infty$ then $(\sum a_i)(\sum b_j) = \sum a_i b_j$, where the sum on the right is evaluated by any method of summation.*

4. Differentiation of power series. It is easy to show that a power series convergent for some $z \neq 0$ has a circle of convergence, at interior points of which it converges absolutely. If we can only justify the instinctive method of differentiating such series, the road is clear not only for generalizing the "elementary

functions," but for developing a good part of "elementary function theory."

Unfortunately a direct attack on the problem of differentiation leads to analytic complexities that are entirely out of proportion to the depth of the problem. In most treatises the proof is measured by paragraphs and pages, rather than lines.

This melancholy fact is all the more distressing because the differentiation at $z=0$ happens to be trivial. If $g(z)=\sum b_n z^n$ for $|z|<r$ then, by inspection,

$$h^{-1}[g(h)-g(0)] = b_1 + O(|h|) \to b_1, \tag{3b}$$

and hence $g'(0)=b_1$.

COROLLARY 4. *When z is interior to the circle of convergence of $f(z)=\sum a_n z^n$ then $f'(z)=\sum n a_n z^{n-1}$.*

For proof let h be so small that $|z|+|h|$ is interior to the circle of convergence. The series

$$\sum |a_n| (|z| + |h|)^n = \sum |a_n| C_j^n |z|^j |h|^{n-j}$$

then converges, C_j^n being the binomial coefficients. The latter series is summed on n and j, but it is easy to imagine the terms enumerated in a definite order, giving a single series. Using this same order, we form the corresponding series without absolute values,

$$\sum a_n C_j^n z^j h^{n-j}. \tag{4}$$

Now collect the terms with $a_0, a_1, a_2, \cdots$ as coefficients on the one hand, and collect the terms with $1, h, h^2, \cdots$ as coefficients on the other. Since the two series so obtained are both scrambled forms of the absolutely convergent series (4), Theorem 1 asserts that they are equal. That is,

$$f(z+h) = \sum a_n(z+h)^n = \sum b_n h^n = g(h),$$

where, in particular, $b_1=a_1+2a_2z+3a_3z^2+\cdots$. Using (3b) we get the desired result,

$$h^{-1}[f(z+h)-f(z)] = h^{-1}[g(h)-g(0)] \to b_1.$$

5. Composite functions. The foregoing proof establishes the validity of the formal expansion $f(z+h)=g(h)$ and hence, it suggests a broader context. Instead of the change of variable $h\to z+h$ we could consider $h\to w(h)$, so that the desired equality is $f[w(h)]=\sum b_n h^n$. Writing z in place of h we shall establish:

COROLLARY 5. *Let $f(z)=\sum a_n z^n$, let $w(z)=\sum w_n z^n$, and let the formal expansion of $f[w(z)]$ be $\sum b_n z^n = g(z)$. If $|z|$ is so small that $\sum |w_n| |z|^n$ is interior to the circle of convergence for f, then $f(w)=g(z)$.*

By formal expansion

$$\sum |a_n| [\sum |w_k| |z|^k]^n = \sum |a_n| C_{p\cdots q}^n |w_0|^p \cdots |w_m|^q |z|^m, \tag{5}$$

where $m=p\cdot 0+\cdots+qm$ and $p+\cdots+q=n$. The multinomial coefficients $C^n_{p\cdots q}$ are positive. The right expression is a complicated sum on $n, p, \cdots, q, m$, but we can still imagine the terms are enumerated in some definite order, giving a single series. The left-hand series is a scrambled form of the right. By hypotheses the left-hand series converges. Hence by Theorem 1, the right-hand series converges.

Corresponding to the right-hand side of (5) is the series without absolute values,

$$\sum a_n C^n_{p\cdots q} w_s^p \cdots w_m^q z^m. \tag{6}$$

If we collect the coefficients of $a_0, a_1, a_2, \cdots$ on the one hand and the coefficients of $1, z, z^2, \cdots$ on the other, the resulting series are equal, because each of them is a scrambled form of the absolutely convergent series (6). This completes the proof.

6. **An infinite product.** The following useful theorem is given by Titchmarsh in [2]:

COROLLARY 6. *If* $\sum |u_n|$ *converges then* $\prod(1+u_nz)$ *converges and can be rearranged as an absolutely convergent series in powers of* z, *thus*:

$$\prod(1+u_nz) = 1 + z\sum u_n + z^2 \sum_{n\neq m} u_mu_n + \cdots.$$

The proof is similar to that of Corollary 5. For any fixed N it is possible to expand the product

$$\prod_1^N (1+|u_n||z|) \equiv 1 + |u_1||z| + |u_2||z| + |u_1||u_2||z|^2 + \cdots \tag{7}$$

as a finite sum of monomials, each of which is a product of certain $|u_i|$ and a power of $|z|$. Changing N to $N+1$ multiplies the previous sum by $1+|u_{N+1}|\,|z|$, hence merely adds some terms to those we already had. Therefore, the series

$$\sum |u_{n_1}||u_{n_2}| \cdots |u_{n_m}||z|^m \qquad (n_i \neq n_j) \tag{8}$$

obtained for $N\to\infty$ is well defined as a formal expansion. The series (8) converges, because the familiar inequality $1+a\leqq e^a$ shows that the product (7) converges as $N\to\infty$.

Associated with (8) is a similar series (say (8)*) without absolute values. If we collect terms corresponding to successive values of N in

$$\prod_1^N (1+u_nz)$$

on the one hand, and collect terms in $1, z, z^2, \cdots$, on the other, the resulting

series are equal, because they are both scrambled forms of one and the same absolutely convergent series (8)* mentioned above. This completes the proof.

7. Measure in college. We shall be concerned with "intervals," which can be ordinary intervals on the real line, or rectangles in the plane, or rectangular solids in space, or more general sets unlike any of these. But it is assumed that the intersection of two intervals is always an interval, and that each interval I has a measure, $|I|$. The latter is a nonnegative real number satisfying the following two axioms:

AXIOM 1. *If $I_1, I_2, I_3, \cdots$ are disjoint intervals contained in another interval I, then $|I_1|+|I_2|+|I_3|+\cdots \leqq |I|$.*

AXIOM 2. *If I_1, I_2, I_3 are intervals whose union contains another interval I, then $|I_1|+|I_2|+|I_3|+\cdots \geqq |I|$.*

It is not hard to show that these properties hold for the familiar measure of segments, rectangles, and so on, but we shall not pause to give the proof here. Our primary objective is to establish the consistency of measure, as follows:

COROLLARY 7. *Let U be the union of countably many disjoint intervals, $U = I_1 \cup I_2 \cup I_3 \cup \cdots$ and let U be contained in an interval. Then the quantity*

$$|U| = |I_1| + |I_2| + |I_3| + \cdots$$

exists and is independent of the decomposition, $\{I_n\}$.

That $|U|$ exists follows from Axiom 1. If $U = J_1 \cup J_2 \cup J_3 \cup \cdots$ is another decomposition of the indicated type, we can define $K_{mn} = I_m \cap J_n$. The K_{mn} are disjoint intervals whose union is contained in the interval that contains U, and hence by Axiom 1 the series $\sum |K_{mn}|$ is convergent, with any convenient enumeration. Also

$$I_m = I_m \cap U = K_{m1} \cup K_{m2} \cup K_{m3} \cup \cdots$$

and hence, by Axioms 1 and 2,

$$|I_m| = |K_{m1}| + |K_{m2}| + |K_{m3}| + \cdots.$$

Therefore $\sum |I_m|$ is a scrambled form of $\sum |K_{mn}|$. Since the same applies to $\sum |J_m|$, equality follows from Theorem 1.

For broad classes of functions f the region described by

$$a < x < b, \qquad 0 < y < f(x)$$

is representable as union of infinitely many rectangles, all of which have their sides parallel to the axes. Corollary 7 asserts that the area is independent of the mode of decomposition, and thus we are enabled to define the definite integral. The same applies to volumes of the type usually encountered in elementary calculus.

8. Measure in high school. Reasoning *a fortiori*, one would conclude that Corollary 7 automatically settles the question of area and volume in high-school geometry. But actually the latter requires further discussion, for the following reason. Corollary 7 says we will get a unique answer by use of a given family of "intervals," no matter how the decomposition within that family is carried out. But it does not obviously ensure the same answer if we use an entirely different family. (Consider a semicircle decomposed into rectangles perpendicular to the diameter, and then decomposed into other rectangles, making an angle of 60° with the diameter.) This matter must be faced in the schools, even though it can be passed over in silence in the university. The trouble is that the schools do not have a preferred co-ordinate system, and thus, the invariance of measure under rotation cannot be dodged.

The solution is to decompose the new intervals into intervals of the original type, getting a new decomposition of the original area. For example, in the case of the semicircle, the rectangles slanted 60° are themselves unions of rectangles slanted 90°. Similar remarks apply to decomposition into triangles, hexagons, circles, and so on. The matter is summarized by:

COROLLARY 8. *Let $U=U_1\cup U_2\cup U_3\cup \cdots$, where the U_i are disjoint and each of them is a union of countably many disjoint intervals. Let U be contained in some interval I. Then*

$$|U| = |U_1| + |U_2| + |U_3| + \cdots$$

independently of the decomposition $\{U_i\}$.

By hypothesis $U_m=I_{m1}\cup I_{m2}\cup I_{m3}\cdots$, where the I_{mi} are disjoint intervals. The I's for different m's are also disjoint, since the U's are. Thus Corollary 7 gives

$$|U| = \sum |I_{mn}|, \qquad \text{and also } |U_m| = |I_{m1}| + |I_{m2}| + \cdots.$$

Since $\sum |U_m|$ is a scrambled form of $\sum |I_{mn}|$, the result follows.

It should be noted that, apart from a little skill in the algebra of sets and series, the only convergence criterion we have used is the comparison test for positive series. The latter follows from the completeness axiom in the following easily digestible form: Every bounded increasing sequence has a limit.

References

1. Konrad Knopp, Theory and Application of Infinite Series, Hafner, New York, 1948, pp. 136–147, 173–175, 180–181, ff.

2. E. C. Titchmarsh, The Theory of Functions, 2nd ed., Oxford University Press, New York, 1939, p. 34.

PYTHAGOREAN TRIPLES OVER GAUSSIAN DOMAINS

N. E. SEXAUER, University of Utah and Northern Illinois University

1. Introduction. The object of this paper is to present a generalization of the number theoretic solution of the Diophantine equation $x^2+y^2=z^2$ when this equation is considered over a certain type of Gaussian (unique factorization) domain. From specializations of the general theory as herein developed the complete solution of the equation is determined for the integers and the Gaussian integers. The notation and terminology used are those of [1, Chapter IV]; in particular, (a, b) will denote a g.c.d. of the elements a and b and $a \sim b$ will be used if a and b are associates.

2. Preliminary results. In this and the remaining sections A will be a Gaussian domain. The following two propositions will be needed in the sequel. The proofs of these results can be modeled after the proofs of the corresponding results in the domain of integers and will be left to the reader. In the statements of these propositions, a, b, and c will be nonzero elements of A.

PROPOSITION 1. *If $a^2 \sim bc$ and $(b, c) \sim g$, then there exist s, t in A such that $b \sim gs^2$, $c \sim gt^2$ and $(s, t) \sim 1$.*

PROPOSITION 2. *If $b^2 | a^2$, then $b | a$.*

A consequence of this result for which we will have need is: If u_1 is a unit and $a^2 = u_1 b^2$, then there exists a unit u_2 such that $a = u_2 b$ and $u_2^2 = u_1$.

3. The general theory. The concept of a primitive solution will first be developed. In the next two propositions x, y and z will be nonzero elements of A.

PROPOSITION 3. *If $x^2+y^2=z^2$, then $(x, y) \sim (x, z) \sim (y, z) \sim (x, y, z)$.*

Proof. We first show that $(x, y) \sim (x, y, z)$. If $g=(x, y)$, then $g^2 | z^2$ which implies by Proposition 2 that $g | z$, hence $g \sim (g, z) \sim (x, y, z)$. The other conclusions follow in a similar manner.

The proof of the next proposition is straightforward and is left to the reader.

PROPOSITION 4. *If $x^2+y^2=z^2$ and $g \sim (x, y, z)$ so that $x=ag$, $y=bg$ and $z=cg$ for some a, b, c in A, then $a^2+b^2=c^2$ and $1 \sim (a, b, c)$.*

We will consider only those solutions in which all members are nonzero. A primitive solution of the equation $x^2+y^2=z^2$ is now defined to be a solution with the property that $1 \sim (x, y, z)$. Upon noting that x, y, z is a solution if and only if there exists x_1, y_1, z_1, g such that $x=gx_1$, $y=gy_1$, $z=gz_1$ and x_1, y_1, z_1 is a primitive solution, we see that the problem of finding solutions has been reduced to finding primitive solutions. The remainder of this section will be devoted to this problem over Gaussian domains which satisfy the following General Hypotheses:

1. There exists an irreducible element p of A with the property that the quotient ring $A/(p)$ is isomorphic to the Galois field of two elements.

2. The characteristic of A is zero.

Let A satisfy the General Hypotheses. We will say that an element a of A is even with respect to p if a is in the residue class determined by 0, while a will be said to be odd with respect to p if a is in the residue class determined by 1. Since in the sequel we will be concerned with only one fixed irreducible element p, we will drop the phrase "with respect to p" and just say even or odd. It is noted that units are odd, that a is even if and only if there exists q in A such that $a=qp$, and that a is odd if and only if there exists q, u in A with u a unit such that $a=qp+u$. It is also of interest to note that if even and oddness are defined by the above characterizations of these concepts, then General Hypothesis 1 is equivalent to the two conditions: (i) There exists an irreducible element p of A such that every element of A is either even or odd. (ii) If u_1, u_2 are units, then $u_1 \pm u_2$ is even. It is also easy to see that all the elementary properties of even and oddness hold including the fact that if $(a, b) \sim 1$, then either a or b is odd. Finally General Hypothesis 2 assures us that the integers are embedded in A, hence since $1+1$ is even we see that $p \mid 2$ in A and therefore the even integers are even with respect to p and the odd integers are odd with respect to p.

Since $p \mid 2$ in A, 2 has the form $2=vp^n p_1^{n_1} \cdots p_m^{n_m}$ where $n \geqq 1$, $n_i \geqq 1$, v a unit, p_i irreducible and not associate to p or p_j for $j \neq i$. By this notation we do not mean to exclude the case $2=vp^n$. The symbols $v, p, p_1, \cdots, p_m, n, n_1, \cdots, n_m$ will be reserved in this context for the rest of this section. As will be seen, it is possible that the equation $x^2+y^2=z^2$ has primitive solutions with x and y of opposite parity or both odd. Two theorems are now given which classify the primitive solutions in these two cases. Note that the ordered triple (x, y, z) is a primitive solution if and only if (y, x, z) is a primitive solution; therefore there will be no loss of generality if we restrict ourselves, as we now do, to the study of primitive solutions (x, y, z) in which x is odd.

In the statements of the three theorems of this section A will satisfy the General Hypotheses. If b is a nonzero factor of a we use the notation a/b to denote the unique q in A such that $a=bq$. We also let $P=2/vp=dD$ where d and D are complementary divisors, in particular $d=p^k p_1^{k_1} \cdots p_m^{k_m}$ where $0 \leqq k \leqq n-1$, $0 \leqq k_i \leqq n_i$ $(i=1, 2, \cdots, m)$. We see that the number of (nonassociated) choices for d is $N=n(n_1+1) \cdots (n_m+1)$.

THEOREM 1. *All primitive solutions in A of the equation $x^2+y^2=z^2$ with x odd and y even are given (perhaps with repetition) by $x=u_3(s^2-u_2t^2)/D$, $y=u_1pdst$, $z=u_3(s^2+u_2t^2)/D$, where u_1, u_2, u_3 are units with the property that $u_2=(u_1/u_3v)^2$ and s and t are parameters such that $(s, t) \sim 1$, x and z are in A, z is odd and each irreducible factor of d is not a factor of z.*

Proof. If x, y, z have the indicated form, it is easy to verify that $D^2(z^2-x^2) = D^2y^2$, so that x, y, z is a solution. Since z is odd and y is even, it follows that x is odd. If we suppose $(y, z) \nsim 1$, there exists an irreducible r such that $r \mid y$ and $r \mid z$. Therefore r is not a factor of d and $r \nsim p$, since z is odd. Since $y=u_1pdst$ it

follows that $r|s$ (or $r|t$); then since $r|z$ we have that $r|t$ (or $r|s$). Hence $(s, t)\not\sim 1$, a contradiction. Thus $(y, z)\sim 1$ and the solution is primitive.

Conversely, let x, y, z be a primitive solution with x odd and y even. Then z is odd, so there exist q_1 and q_2 such that $z+x=q_1p$, $z-x=q_2p$. From $(q_1+q_2)p=2z=vpPz$ and $(q_1-q_2)p=2x=vpPx$, we have $Px=v^{-1}(q_1-q_2)$ and $Pz=v^{-1}(q_1+q_2)$. If $(q_1, q_2)\not\sim 1$, let r be an irreducible factor of (q_1, q_2). Since $(x, z)\sim 1$, it follows that $r|P$. Hence $(q_1, q_2)\sim p^jp_1^{j_1}\cdots p_m^{j_m}$ for certain nonnegative integers j, $j_1, \cdots, j_m$. There exist elements h_1 and h_2 in A such that $Px=h_1(q_1, q_2)$, $Pz=h_2(q_1, q_2)$. If $j>n-1$, it follows from the last two equations that $p|(x, z)$, a contradiction; hence $0\leqq j\leqq n-1$. Similarly $0\leqq j_i\leqq n_i$. Hence $(q_1, q_2)\sim d$, one of the N divisors of P. Since y is even, $y=qp$. Then $y^2=q^2p^2=z^2-x^2=q_1q_2p^2$ implies $q^2=q_1q_2$. By Proposition 1 there exist s, t with $(s, t)\sim 1$ and units v_1, v_2 such that $q_1=v_1s^2d$, $q_2=v_2t^2d$. By the remark following Proposition 2 there exists a unit u_1 so that $u_1^2=v_1v_2$ and $y=u_1pdst$. If we set $u_2=v_1^{-1}v_2$ and $u_3=v^{-1}v_1$, then the relations $Px=v^{-1}(q_1-q_2)$ and $Pz=v^{-1}(q_1+q_2)$ show that x and z have the form stated in the theorem. It is also easy to see that $u_2=(u_1/u_3v)^2$. Finally, if r is an irreducible factor of d, then $r|y$; hence $(y, z)\sim 1$ implies $r\nmid z$.

REMARK 1. In the statement of Theorem 1 the condition that z is odd can be specified as follows:

Case 1. If D is even, the conditions $(s, t)\sim 1$ and $Dz=u_3(s^2+u_2t^2)$ require that s and t must both be odd, say $s=a_1p+1$, $t=a_2p+1$. Let the unit u_2 be expressed by $u_2=a_3p-1$. Then $Dz=u_3(a_1^2p+u_2a_2^2p+a_3)p+u_3(a_1+u_2a_2)vp^2dD$. Since D is even, it follows that z is odd if and only if $(a_1^2p+u_2a_2^2p+a_3)p/D$ is odd.

Case 2. If D is odd, the conditions $Dz=u_3(s^2+u_2t^2)$, $y=u_1pdst$ and $(s, t)\sim 1$ show that z is odd if and only if s and t are nonzero elements of A with opposite parity.

The following notation will be used in Theorems 2 and 3. Let $P_1=2/vp^n=d_1D_1$ where d_1 and D_1 are complementary divisors, in particular $d_1=p_1^{k_1}\cdots p_m^{k_m}$ where $0\leqq k_i\leqq n_i$ $(i=1, 2, \cdots, m)$. We see that the number of (nonassociated) choices for d_1 is $N_1=(n_1+1)\cdots(n_m+1)$.

THEOREM 2. *All primitive solutions in A of the equation $x^2+y^2=z^2$ with x odd and y odd are given (perhaps with repetition) by*

$$x = u_3(s^2 - u_2t^2)/p^nD_1, \qquad y = u_1d_1st, \qquad z = u_3(s^2 + u_2t^2)/p^nD_1$$

where u_1, u_2, u_3 are units with the property that $u_2=(u_1/u_3v)^2$ and s and t are parameters such that s and t are both odd and $(s, t)\sim 1$, x and z are in A, z is even and each irreducible factor of d_1 is not a factor of z.

Proof. If x, y, z have the indicated form, then $p^{2n}D_1^2(z^2-x^2)=p^{2n}D_1^2y^2$ and therefore x, y, z is a solution. Since s and t are both odd, y is odd; hence x is odd since z is even. Using essentially the same proof as employed in the proof of Theorem 1 we can show that $(y, z)\sim 1$.

Conversely, let x, y, z be a primitive solution with x odd and y odd. Then z is even and therefore $z+x=q_1$ and $z-x=q_2$ are both odd. From $q_1+q_2=2z=vp^nP_1z$ and $q_1-q_2=2x=vp^nP_1x$, we have $p^nP_1z=v^{-1}(q_1+q_2)$ and $p^nP_1x=v^{-1}(q_1-q_2)$. It follows as in the proof of Theorem 1 that $(q_1, q_2)\sim d_1$, one of the N_1 divisors of P_1; the only modification in the proof is that r cannot be an associate of p since q_1 is odd. Since $y^2=z^2-x^2=q_1q_2$ we see by use of Proposition 1 that there exist s, t with $(s, t)\sim 1$ and units v_1, v_2 such that $q_1=v_1s^2d_1$, $q_2=v_2t^2d_1$; from which it follows that s and t are both odd. Finally using the same proof as employed in Theorem 1 we can show that there exist units u_1, u_2, u_3 such that x, y, z have the form stated in the theorem, $u_2=(u_1/u_3v)^2$ and that each irreducible factor of d_1 is not a factor of z.

REMARK 2. In the statement of Theorem 2 the condition that z is even can be characterized as follows: Let s, t and u_2 be expressed by $s=a_1p+1$, $t=a_2p+1$ and $u_2=a_3p-1$. Then

$$p^nD_1z = u_3(a_1^2p + u_2a_2^2p + a_3)p + u_3(a_1 + u_2a_2)vd_1D_1p^{n+1}.$$

It follows that $p^{n-1}D_1$ is a factor of $u_3(a_1^2p+u_2a_2^2p+a_3)$ and that z is even if and only if $(a_1^2p+u_2a_2^2p+a_3)/p^{n-1}D_1$ is even.

Since the integers are embedded in A, primitive solutions in which x is odd and y is even always exist. In the next section we will see that there exist domains in which there are primitive solutions in which x and y are both odd and also domains in which such solutions do not exist. Therefore a natural question is: Is there a criterion under which we can determine whether these "odd-odd" solutions do or do not exist? An answer to this question is given in the next theorem.

THEOREM 3. *The equation $x^2+y^2=z^2$ has primitive solutions in A in which x and y are both odd if and only if n is an even integer.*

Proof. Let $n=2n'$. Take a_1, a_2 in A with the property that a_1 and a_2 have opposite parity and let $s=a_1p^{n'}+1$ and $t=a_2p^{n'}+1$. Then $s^2-t^2=p^nx$ where $x=a_1^2-a_2^2+vP_1(a_1-a_2)p^{n'}$ and $s^2+t^2=p^nz$, where $z=a_1^2+a_2^2+vP_1(a_1+a_2)p^{n'}+vP_1$. Since a_1 and a_2 have opposite parity it is seen from these expressions for x and z that x is odd and z is even. If $y=vP_1st$, then y is odd. Since $p^{2n}(z^2-x^2)=p^{2n}y^2$ we see that x, y, z is a solution. If $g\sim(x, y, z)$ and $x=x_1g$, $y=y_1g$ and $z=z_1g$, then x_1, y_1, z_1 is a primitive solution in which x_1 and y_1 are both odd.

Conversely let $n=2n'+1$. Assume by way of contradiction that $x=bp+1$, $y=cp+1$, $z=ap$ is a solution. Upon substituting these expressions into the equation and cancelling p we have that $a^2p=(b^2+c^2)p+ep^{2n'+1}+vP_1p^{2n'}$ where $e=(b+c)vP_1$. If $n'=0$, the left side of this equation is even while the right side is odd. In the case when $n'\geqq 1$ we use a "descent" argument to obtain a contradiction by showing that for each integer k such that $0\leqq k\leqq n'$, there exist a, b, c, e in A such that $a^2p=(b^2+c^2)p+ep^{2k+1}+vP_1p^{2k}$. To this end assume that $0<k+1\leqq n'$ and that there are elements of A such that $a^2p=(b^2+c^2)p+ep^{2k+3}+vP_1p^{2k+2}$. If a is odd, then upon cancelling p from the given equation we see

that b and c have opposite parity. Let $a=a_1p+1$ and without loss of generality let $b=b_1p$ and $c=c_1p+1$. It follows that

$$a_1^2p = (b_1^2+c_1^2)p + (e + c_1vP_1p^{2n'-2k} - a_1vP_1p^{2n'-2k})p^{2k+1} + vP_1p^{2k}$$

which shows that the statement holds for the integer k. A similar result is obtained if a is even since in this case b and c must have the same parity. Finally we see that when we use the above statement with $k=0$ we obtain the same contradiction as in the case when $n'=0$.

4. Two applications. Recall that we need to consider only primitive solutions in which x is odd, and that in Theorems 1 and 2 for each choice of d and d_1 respectively we obtain a particular form of the solutions.

THEOREM 4. *The totality of primitive solutions of $x^2+y^2=z^2$ in the domain of integers is given by the following two formulas where s and t range over the nonzero integers subject to the conditions that $(s,t)\sim 1$ and that s and t have opposite parity with respect to* 2.

$$1.\quad \begin{aligned} x &= s^2-t^2\\ y &= 2st\\ z &= s^2+t^2 \end{aligned} \qquad\qquad 2.\quad \begin{aligned} x &= -(s^2-t^2)\\ y &= -2st\\ z &= -(s^2+t^2). \end{aligned}$$

Proof. If we let $p=2$, then the General Hypotheses are satisfied and in the factorization of 2 the exponent of p is 1. Thus the single form given by Theorem 1 (corresponding to $d=1=D$) will give us all the primitive solutions. The combinations of the units that satisfy the unit condition $u_2=(u_1/u_3v)^2$ where $v=1$ are $u_2=1$, $u_1=\pm 1=u_3$. Therefore upon using the result given in Case 2 of Remark 1 we see that the four formulas

$$1.\ \begin{aligned} x &= s^2-t^2\\ y &= 2st\\ z &= s^2+t^2 \end{aligned} \qquad 2.\ \begin{aligned} x &= -(s^2-t^2)\\ y &= 2st\\ z &= -(s^2+t^2) \end{aligned} \qquad 3.\ \begin{aligned} x &= s^2-t^2\\ y &= -2st\\ z &= s^2+t^2 \end{aligned} \qquad 4.\ \begin{aligned} x &= -(s^2-t^2)\\ y &= -2st\\ z &= -(s^2+t^2) \end{aligned}$$

will give us the primitive solutions as s and t range over the nonzero integers subject to the conditions as stated in the conclusion of this theorem. If s is replaced by $-s$ and t by t in formula 1, then formula 3 is obtained. From this it is easy to see that formulas 1 and 3 are equivalent in the sense that each formula gives the same set of solutions of the equation. Similarly formulas 2 and 4 are equivalent. It is clear that formulas 1 and 4 are independent, that is, not equivalent.

THEOREM 5. *The totality of primitive solutions of $x^2+y^2=z^2$ in the domain G of Gaussian integers is given by the following set of four independent formulas where in* 1. *and* 2. *s and t range over the nonzero elements of G subject to the conditions that $(s,t)\sim 1$ and that they have opposite parity with respect to $1+i$, and where in* 3. *and* 4., *$s=a_1(1+i)+1$, $t=a_2(1+i)+1$ and a_1 and a_2 range over G*

subject to the conditions that a_1 and a_2 have opposite parity with respect to $1+i$ and that $(s, t)\sim 1$.

$x = s^2 - t^2$	$x = i(s^2 - t^2)$	$x = (s^2 - t^2)/2$	$x = i(s^2 - t^2)/2$
1. $y = 2st$	2. $y = 2ist$	3. $y = st$	4. $y = ist$
$z = s^2 + t^2$	$z = i(s^2 + t^2)$	$z = (s^2 + t^2)/2$	$z = i(s^2 + t^2)/2.$

Proof. Under the function $\delta(a+bi)=a^2+b^2$, G is a Euclidean domain [1, p. 123]. If $p=1+i$, then p is irreducible [2, p. 17] and since $\delta(p)=2$, if we use the division algorithm to divide an element of G by p then $\delta(r)$, where r is the remainder, must be 0 or 1, that is r is 0 or a unit. Since the units of G are ± 1 and $\pm i$ [2, p. 7], one can easily verify that $u_1 \pm u_2$ is even as u_1, u_2 range over the units. Thus since $2=(-i)p^2$ the primitive solutions with x odd and y even with respect to $1+i$ are of one of the two forms as listed in Theorem 1. Note that in the first form (corresponding to $d=1$, $D=p$) the condition that z is odd is equivalent to: if $1+u_2=a_3p$, then a_3 is odd; this follows by use of the result given in Case 1 of Remark 1. But there are no solutions of the first form since there are no units u_1, u_2, u_3 in G such that $1+u_2=a_3p$ with a_3 odd and using $v=-i$ such that $u_2=-(u_1/u_3)^2$. It can be seen that there are exactly 16 combinations of the units of G that satisfy the unit condition $u_3^2u_2=-u_1^2$. Upon substituting each of these combinations into the formulas for x, y and z given in the second form of Theorem 1 (corresponding to $d=p$, $D=1$) and using the result given in Case 2 of Remark 1, we obtain 16 formulas for primitive solutions, two of which are 1 and 2 as listed in the conclusion of this theorem. Techniques similar to those employed in the proof of Theorem 4 will show that all but 1 and 2 can be eliminated.

The primitive solutions in which both x and y are odd are given by the single form (corresponding to $d_1=1=D_1$) of Theorem 2. It is seen from the unit condition $u_2=-(u_1/u_3)^2$ that u_2 must be either $+1$ or -1. If $u_2=1$, then the condition "$(a_1^2p+u_2a_2^2p+a_3)/p$ is even" given in Remark 2 is equivalent to "a_1 and a_2 have opposite parity," while if $u_2=-1$, then this same condition is equivalent to "a_1 and a_2 have the same parity." When $u_2=1$, then $u_1^2=-u_3^2$ and if we substitute the eight combinations of units that satisfy this condition into the general form given in Theorem 2 we obtain eight formulas for primitive solutions two of which are 3 and 4 as listed in the conclusion of this theorem. When $u_2=-1$, then $u_1^2=u_3^2$ and we obtain eight additional formulas, but in these a_1 and a_2 will have the same parity. As above all but 3 and 4 can be eliminated. Finally, it can be shown that the four listed formulas are independent.

The author wishes to thank the referee for helpful criticisms and suggestions.

References

1. N. Jacobson, Lectures in Abstract Algebra, Vol. I, Van Nostrand, Princeton, N. J., 1951.

2. H. Pollard, The theory of algebraic numbers, Carus Mathematical Monograph Number 9, MAA, 1950.

SOME REMARKS ON INVERTIBLE SPACES

YIM-MING WONG, University of Hong Kong

1. Introduction. Invertible spaces were first defined by P. H. Doyle and J. G. Hocking in [1] as follows:

DEFINITION 1.1. *A topological space X is said to be completely invertible if, for each nonempty open subset G of X, there is an onto homeomorphism $h: X \to X$ such that $h(CG) \subset G$, where CG denotes the complement of G. h is called an inverting map for G.*

REMARK. To distinguish our definition (see Chapter 2.1 below) from that of Doyle and Hocking's, we shall, throughout this paper, call a space which is invertible in the sense of Doyle and Hocking a completely invertible space, as has been done in the above definition.

Using diverse methods, the authors of [1], [2] and [3] have proved among others the following results:

THEOREM 1.1. *Let X be a completely invertible space. If there is a non-empty open subset G of X having one of the following properties*:

> (1) T_0, (2) T_1, (3) *separability*, (4) *satisfying the first axiom of countability*, (5) *satisfying the second axiom of countability*, (6) *having a compact closure*, (7) *being Hausdorff*, (8) *being regular*, (9) *being normal*, (10) *metrizability*,

then X has also the corresponding property.

In this paper, we improve these results by weakening the condition and providing simpler proofs.

2. T_0, T_1, separability, satisfying the first axiom of countability, satisfying the second axiom of countability, having a compact closure and being Lindelöf.

LEMMA 2.1. *Let A, B be two open sets in a topological space X such that $X = A \cup B$. If A and B have one of the properties* (1)–(6), *stated in Section* 1, *or are Lindelöf, then X has also the corresponding property.*

Proof. Let A, B be T_0, and $x, y \in X$ such that $x \neq y$. We now prove for x or y that there exists an open neighborhood in X which does not contain the other. If x, y are not at the same time in A or B, then the existence of such an open neighborhood is obvious, (namely, A or B). If x, y are at the same time in A, then since A is T_0, there is an open neighborhood U of x or y in A, and also in X as A is open, which does not contain the other. It is similar for the case that x, y are in B. We have then proved in any case the existence of such an open neighborhood, hence X is T_0.

We can prove similarly that if A, B are T_1 then X is also T_1.

Let A, B be separable and A_1, B_1 be countable dense subsets of A, B respec-

tively. We see first that $A_1 \cup B_1$ is countable. By a theorem stated in [4, p. 53], since $A \setminus B$ and $B \setminus A$ are separated,

$$(A_1 \cup B_1)^- = ((A_1 \cap A)^- \cap A) \cup ((B_1 \cap B)^- \cap B) = (A_1^- \cap A) \cup (B_1^- \cap B)$$
$$= A \cup B = X.$$

This proves that $A_1 \cup B_1$ is dense in X; hence X is separable.

In case A, B have one of properties (4)–(6) and are Lindelöf, the verification that X has the corresponding property is straightforward.

DEFINITION 2.1. *A topological space X is said to be invertible with respect to an open subset G of itself if there is an onto homeomorphism $h: X \to X$ such that $h(CG) \subset G$. This homeomorphism h is called an inverting map for G.*

REMARK. Obviously, if X is invertible with respect to every nonempty open subset of itself then it is completely invertible. However, a topological space X which is invertible with respect to an open subset need not be completely invertible. For example, let $R = \{x: -\infty < x < +\infty\}$, the real line with the usual topology, let $G = \{x: x < 1\}$ and let h be the homeomorphism $R \to R$ defined by $h(x) = -x$. Then $h(CG) \subset G$ so that R is invertible with respect to the open subset G. But X is not completely invertible. In fact, by Theorem 1.1, if a completely invertible space is locally compact then it is compact. Since R is locally compact but not compact, R is not completely invertible.

From the above remark we shall see that the following theorem is stronger than that stated in Section 1.

THEOREM 2.1. *Let X be a topological space invertible with respect to an open subset G. If G has one of the properties* (1)–(6), *stated in Section* 1, *or is Lindelöf, then X has also the corresponding property.*

Proof. Let h be an inverting map for G. Then h^{-1} is also a homeomorphism of X onto itself and $h^{-1}(G)$ is an open subset of X. Since $h(CG) \subset G$, we have $h^{-1}(G) \supset CG$, and hence $X = G \cup h^{-1}(G)$.

Now all the properties mentioned in Theorem 2.1 are topological; therefore, if G has one of these properties, then $h^{-1}(G)$ has also this property.

It is shown in Lemma 2.1, however, that if X is the union of two open subsets having one of the above mentioned properties, then X itself also has the corresponding property.

The theorem is thus proved.

3. Being Hausdorff, regular or normal. In what follows, for two subsets A, B of a topological space X, we shall say that they are strongly separated if there are two open sets U, V such that $A \subset U$, $B \subset V$ and $U \cap V = \emptyset$.

DEFINITION 3.1. *Let X be a topological space, G_1, G_2 two open subsets of X such that $X = G_1 \cup G_2$.*

(i) *If for each $x \in G_1 \setminus G_2$ and $y \in G_2 \setminus G_1$, $\{x\}$ and $\{y\}$ are strongly separated, then X is said to have s_0-property with respect to G_1 and G_2.*

(ii) *If for each* $x \in G_2 \backslash G_1$, $\{x\}$ *and* $G_1 \backslash G_2$ *are strongly separated, and for each* $y \in G_1 \backslash G_2$, $\{y\}$ *and* $G_2 \backslash G_1$ *are strongly separated, then* X *is said to have* s_1*-property with respect to* G_1 *and* G_2.

(iii) *If* $G_1 \backslash G_2$ *and* $G_2 \backslash G_1$ *are strongly separated, then* X *is said to have* s_2*-property with respect to* G_1 *and* G_2.

REMARK. Obviously, if G_1 and G_2 are open subsets of a topological space X such that $X = G_1 \cup G_2$, then $G_1 \backslash G_2$ and $G_2 \backslash G_1$ are disjoint closed subsets of X. It is then clear that a Hausdorff space has the s_0-property, a regular space has the s_1-property and a normal space has the s_2-property with respect to every pair of such G_1, G_2.

LEMMA 3.1. *Let* X *be a topological space and* A, B *and* C *subsets of* X. *If* A *and* B *are strongly separated and* A *and* C *are strongly separated, then* A *and* $B \cup C$ *are strongly separated.*

Proof. Let U_1 and V_1 be two disjoint open neighborhoods of A and B, and U_2 and V_2 two disjoint neighborhoods of A and C, respectively. It is easy to see that $U = U_1 \cap U_2$ and $V = V_1 \cup V_2$ are disjoint open subsets of X such that $A \subset U$, $B \cup C \subset V$ and $U \cap V = \emptyset$.

LEMMA 3.2. *Let* X *be a topological space and* G *an open subspace of* X. *If two subsets* A, B *of* G *are strongly separated in* G *then they are strongly separated in the space* X.

Proof. It follows from the fact that open subsets of an open subspace are also open in the original space.

LEMMA 3.3. *Let* X, Y *be two topological spaces and* h *a homeomorphism of* X *onto* Y. *Then,* A, B *are strongly separated iff* $h(A)$ *and* $h(B)$ *are strongly separated.*

Proof. This follows from the fact that a homeomorphism is a 1-1 onto mapping and preserves openness.

THEOREM 3.1. *Let* X *be a topological space having the* s_0*-property with respect to* G_1 *and* G_2. *If* G_1 *and* G_2 *are Hausdorff, then* X *is Hausdorff.*

Proof. Let $x, y \in X$ and $x \neq y$. If x, y are at the same time in A then, since A is Hausdorff, $\{x\}$ and $\{y\}$ are strongly separated in A and, by Lemma 3.2, in X also. Similarly, $\{x\}$ and $\{y\}$ are strongly separated when they are at the same time in B. If $x \in A \backslash B$ and $y \in B \backslash A$, then by the s_0-property, they are strongly separated in X.

PROPOSITION 3.1. *Let* X *be a completely invertible space,* G *a nonempty open subset of* X *and* h *an inverting map for* G. *If* G *is Hausdorff, then* X *has the* s_0*-property with respect to* G *and* $h^{-1}(G)$.

Proof. We have seen in the proof of Theorem 2.1 that X is the union of the open sets G and $h^{-1}(G)$. If $A \equiv G \cap h(G) = \emptyset$, then G and $h^{-1}(G)$ are both open and closed, and so X obviously has the s_0-property with respect to G and $h^{-1}(G)$.

If the open set $A \neq \emptyset$, let g be an inverting map for A. It is clear that $g(x)$, $g(y) \in A \subset G$, and $g(x) \neq g(y)$. Since G is Hausdorff they are strongly separated in G, hence by Lemma 3.2 in X also. By Lemma 3.3, $\{x\}$ and $\{y\}$ are strongly separated in X.

COROLLARY 3.1. *Let X be a completely invertible space. If X has a nonempty open subset which is Hausdorff, then X is Hausdorff.*

THEOREM 3.2. *Let X be a topological space having the s_1-property with respect to G_1 and G_2. If G_1 and G_2 are regular, then X is regular.*

Proof. Let F be a closed subset of X and let $x \in X \backslash F$. We assume first that $x \in G_1 \backslash G_2$. Let $F_1 = F \cap G_1$ and $F_2 = F \cap (G_2 \backslash G_1)$. It is seen that F_1 is closed in G_1, F_2 is closed in both $G_2 \backslash G_1$ and X, and $F = F_1 \cup F_2$. On the one hand since G_1 is regular and open, by Lemma 3.2, $\{x\}$ and F_1 are strongly separated. On the other hand, since X has the s_1-property with respect to G_1 and G_2, $\{x\}$ and F_2 are strongly separated. By Lemma 3.1, since $F = F_1 \cup F_2$, $\{x\}$ and F are strongly separated. The proof is similar in case $x \in G_2 \backslash G_1$. If $x \in G_1 \cap G_2$ then, by Lemma 3.2, since G_1 and G_2 are regular and open, $\{x\}$ and $F \cap G_i$ are separated for $i = 1, 2$. By Lemma 3.1, $\{x\}$ and F are strongly separated.

The theorem is thus proved.

PROPOSITION 3.2. *Let X be a completely invertible space, G a nonempty open subset of X, and h an inverting map for G. If G is regular, then X has the s_1-property with respect to G and $h^{-1}(G)$.*

Proof. We have seen in the proof of Theorem 2.1 that X is the union of the open sets G and $h^{-1}(G)$. If $A \equiv G \cap h^{-1}(G) = \emptyset$, then G and $h^{-1}(G)$ are both open and closed, and so X obviously has the s_1-property with respect to G and $h^{-1}(G)$.

If $A \neq \emptyset$, let $F_1 = G \backslash h^{-1}(G)$ and $F_2 = h^{-1}(G) \backslash G$, and let g be an inverting map for A. It is clear that $g(F_1)$ and $g(F_2)$ are disjoint closed subsets of X both contained in A, and hence in G also. Since by assumption G is regular, for each $x \in F_1$, $\{g(x)\}$ and $g(F_2)$ are strongly separated in G and, by Lemma 3.2, in X also. By Lemma 3.3, $\{x\}$ and F_2 are strongly separated.

We can prove similarly that, for each $y \in F_2$, $\{y\}$ and F_1 are strongly separated.

This proves that X has s_1-property with respect to G and $h^{-1}(G)$.

COROLLARY 3.2. *Let X be a completely invertible space. If X has a nonempty open subset which is regular, then X is regular.*

THEOREM 3.3. *Let X be a topological space having the s_2-property with respect to G_1 and G_2. If G_1 and G_2 are normal, then X is normal.*

Proof. Let S and T be two disjoint closed subsets of X. Consider the following subsets:

$$S_1 = S \cap (G_1 \backslash G_2); \qquad S_2 = S \cap (G_1 \cap G_2); \qquad S_3 = S \cap (G_2 \backslash G_1);$$

$$T_1 = T \cap (G_1 \backslash G_2); \qquad T_2 = T \cap (G_1 \cap G_2); \qquad T_3 = T \cap (G_2 \backslash G_1).$$

It is seen that $S=S_1\cup S_2\cup S_3$ and $T=T_1\cup T_2\cup T_3$.

On the one hand, by assumption, G_1 is an open and normal subspace of X. Since $S_1\cup S_2$, $T_1\cup T_2$ are disjoint closed subsets of G_1, they are strongly separated in X. It is then clear that

(i) T_1 and $S_1\cup S_2$ are strongly separated;

(ii) T_2 and $S_1\cup S_2$ are strongly separated.

Similarly, we see also that

(iii) T_3 and $S_2\cup S_3$ are strongly separated;

(iv) T_2 and $S_2\cup S_3$ are strongly separated.

On the other hand, since by assumption X has the s_2-property with respect to G_1 and G_2, we see that

(v) T_1 and S_3 are strongly separated;

(vi) T_3 and S_1 are strongly separated.

By making use of Lemma 3.1, we obtain from (i)–(vi):

(a) T_1 and S are strongly separated;

(b) T_2 and S are strongly separated;

(c) T_3 and S are strongly separated.

Using Lemma 3.1 again, we see from (a)–(c) that T and S are strongly separated.

The theorem is thus proved.

PROPOSITION 3.3. *Let X be a completely invertible space, G a nonempty open subset of X, and h an inverting map for G. If G is normal, then X has the s_2-property with respect to G and $h^{-1}(G)$.*

Proof. It has been seen in the proof of Theorem 2.1 that G and $h^{-1}(G)$ are both open in X, and $X=G\cup h^{-1}(G)$. If $G\cap h^{-1}(G)=\emptyset$, then the assertion is trivial. Let $A\equiv G\cap h^{-1}(G)\neq\emptyset$, and let g be an inverting map for A. Then $g(G\backslash h^{-1}(G))$ and $g(h^{-1}(G)\backslash G)$ are closed (in X) disjoint subsets in A, and hence in G also. Since G is normal by assumption, $g(G\backslash h^{-1}(G))$ and $g(h^{-1}(G)\backslash G)$ are strongly separated in G and, by Lemma 3.2, in X also. By Lemma 3.3, $G\backslash h^{-1}(G)$ and $h^{-1}(G)\backslash g$ are strongly separated. Thus X has the s_2-property with respect to G and $h^{-1}(G)$, as was to be proved.

COROLLARY 3.3. *Let X be a completely invertible space. If X has a nonempty open subset which is normal, then X is normal.*

REMARK. As has been pointed out in the Remark after Definition 2.1, the real line R with the usual topology is not completely invertible. But it is easy to see that R has the s_0-property, s_1-property and s_2-property with respect to $G_1=\{x: x<1\}$ and $G_2=\{x: x>-1\}$. This shows that our Theorems 3.1, 3.2 and 3.3 are stronger than the corresponding results stated in Section 1.

4. Metrizability.

LEMMA 4.1. *Let X be a topological space and S a closed subspace of X. If B, a family of subsets of S, is σ-locally finite in S, then it is σ-locally finite in X.*

Proof. Let $B=\bigcup_{i=1}^{\infty} B_i$, such that B_i is locally finite in S for each i. For $x\in X$,

if $x \in S$ then there is an open subset G of S such that $x \in G$ and G intersects only finitely many members of B_i. Being a relative open set G can be written as $G = G_1 \cap S$ where G_1 is open in X. It is clear that G_1 intersects also only finitely many members of B_i. If $x \in CS$, the complement of S, then, since S is closed, CS is an open set containing x which does not intersect any member of B_i. Therefore B_i is locally finite in X, and hence B is σ-locally finite in X.

THEOREM 4.1. *Let X be a topological space, invertible with respect to an open subset G with inverting map h. If X has the s_1-property with respect to G and $h^{-1}(G)$, and if G^- is metrizable, then X is metrizable.*

Proof. By a metrization theorem in [**4**, p. 127], we know that a topological space X is metrizable iff it is T_1, regular, and its topology has a σ-locally finite base.

Let X satisfy the assumptions stated in Theorem 4.1. Then by Theorem 2.1 and Theorem 3.2, X is T_1 and regular. It remains to show that the topology of X has a σ-locally finite base. Since G^- is by assumption metrizable, it has a σ-locally finite base, say B_1. But having a σ-locally finite base is a topological property; therefore $h^{-1}(G^-)$ has a σ-locally finite base, say B_2.

By Lemma 4.1, B_1 and B_2 are σ-locally finite in X; then so is $B_1 \cup B_2$. Let $\tilde{B}_1 = \{B \cap G: B \in B_2\}$ and $\tilde{B}_2 = \{B \cap h^{-1}(G): B \in B_2\}$. It is clear that $\tilde{B}_1$ and $\tilde{B}_2$ are bases for G and $h^{-1}(G)$ respectively. Hence $\tilde{B}_1 \cup \tilde{B}_2$ is a base for X, as each open set O in X is the union of $O \cap G$ and $O \cap h^{-1}(G)$. Since $B_1 \cup B_2$ is σ-locally finite in X, $\tilde{B}_1 \cup \tilde{B}_2$ is also σ-locally finite in X. Thus, X is metrizable, as was to be proved.

REMARK. If X is metrizable, then obviously X has the s_1-property with respect to G and $h^{-1}(G)$, and G^- is metrizable.

5. An example. The following example shows that for being Hausdorff, regular, normal, and metrizable, a theorem of the same type as Theorem 2.1 does not hold.

Example. Let τ be the usual topology for the real line R. Consider the following family of subsets of R:

$$\tilde{\tau} = \{A: A \in \tau \text{ and } 1, -1 \in A\} \cup \{A \backslash \{1\}: A \in \tau,$$
$$1 \in A\} \cup \{A \backslash \{-1\}: A \in \tau, -1 \in A\}.$$

It is easy to verify that $(R, \tilde{\tau})$ is a topological space. Let $G = \{x: x \in R$ and $x \neq 1\}$ and the homeomorphism $h: R \rightarrow R$ defined by $h(x) = -x$. It is easily seen that $(R, \tilde{\tau})$ is invertible with respect to G with the inverting map h.

We now prove that G is Hausdorff, regular, normal and metrizable, but $(R, \tilde{\tau})$ is not.

Since a metrizable topological space is Hausdorff, regular and normal we need only to prove that G is metrizable but $(R, \tilde{\tau})$ is neither Hausdorff nor regular nor normal.

Consider the map $f\colon R \to R \times R$ defined by

$$f(t) = (t^2, (t-1)(t+1)).$$

It is easy to verify that the restriction of f to G, as a subspace of $(R, \tilde{\tau})$, is a homeomorphism of G onto the subspace $f(G)$ of $R \times R$, with the usual metric topology. Since $f(G)$ is a metric space, G is metrizable.

Since $\{1\}$ and $\{-1\}$ are disjoint closed subsets of $(R, \tilde{\tau})$ which are not strongly separated, $(R, \tilde{\tau})$ is neither Hausdorff nor regular nor normal.

The author wishes to thank Professor Young-Chow Wong for his supervision during the preparation of this paper.

References

1. P. H. Doyle and J. G. Hocking, Invertible spaces, this MONTHLY, 68(1961) 959–965.
2. L. Norman, Some remarks on invertible spaces, this MONTHLY, 70(1963) 181–183.
3. W. J. Gray, On the metrizability of invertible spaces, this MONTHLY, 71(1964) 533–534.
4. J. L. Kelley, General Topology, Van Nostrand, Princeton, N. J., 1955.

SHRINKING BASES IN BANACH SPACES

J. R. RETHERFORD, Louisiana State University

1. Introduction. A Schauder basis in a Banach space X is a sequence $\{x_i\}$ with the property that for each $x \in X$ there is a unique sequence of scalars $\{a_i\}$ such that $x = \sum_1^\infty a_i x_i$, convergence in the norm topology. If f_i is defined by $f_i(x) = a_i$ then f_i is a continuous linear functional on X and $f_i(x_j) = \delta_{ij}$, the Kronecker delta. Throughout this note we will write $\{x_i, f_i\}$ for the basis $\{x_i\}$ in X and the associated sequence of coefficient functionals $\{f_i\}$ in X^* and speak of the basis $\{x_i, f_i\}$ for X.

A basis $\{x_i, f_i\}$ for X is

(A) *shrinking* if $\lim_n \|f\|_n = 0$, where $\|f\|_n$ is the norm of $f \in X^*$ on the closed linear span of $\{x_i\}_{i=n}^\infty$.

This notion is due to James [4], although the word shrinking appears to have been introduced by Day [2]. The usefulness of this concept has been aptly demonstrated by James [4, 5, 6] with his now famous characterization of reflexive spaces with bases which led to the construction of a nonreflexive Banach space isometrically isomorphic to its second conjugate space.

In this expository paper we introduce a number of concepts which are shown to be equivalent to shrinking. While most of these equivalences are

known, they are scattered throughout the literature and this appears to be the first offering of a "cyclic" proof of these equivalences.

Included in the bibliography is every paper, known to the author, concerning shrinking bases.

2. Notation and facts concerning bases. If $\{y_i\}$ is a sequence in a Banach space Y, $[y_i | i \in \omega]$, $[y_i | i \leqq n]$ and $[y_i | i \geqq n]$ denote the closed linear spans of $\{y_i\}_{i=1}^{\infty}$, $\{y_i\}_{i=1}^{n}$ and $\{y_i\}_{i=n}^{\infty}$ respectively.

A sequence $\{y_n\}$ in a Banach space Y is said to be a *basic sequence in* Y if $\{y_n\}$ is a Schauder basis for $[y_n | n \in \omega]$. We say that $\{y_n, g_n\}$ is *basic in* Y if $\{y_n\}$ is a basic sequence in Y and $\{g_n\}$ is the associated sequence of coefficient functionals.

The following well-known results are stated here for easy reference.

THEOREM 2.1. *If $\{x_i, f_i\}$ is a basis for a Banach space X then*

(i) *$\{f_i, Qx_i\}$ is basic in X^*, where Q is the canonical map of X into X^{**};*

(ii) *if $U_n(x) = \sum_1^n f_i(x)x_i$ and $V_n(x) = \sum_{n+1}^{\infty} f_i(x)x_i$ then there is a $K > 0$ such that $\|U_n\| \leqq K$ and $\|V_n\| \leqq K$ for each $n \in \omega$;*

(iii) *for each $F \in X^{**}$, $\sup_n \|\sum_1^n F(f_i)x_i\| < +\infty$; if $\{a_n\}$ is such that $\sup_n \|\sum_1^n a_i x_i\| < +\infty$ then there is an $F \in X^{**}$ with $F(f_i) = a_i$;*

(iv) *for each $h \in X^*$, $\sup_n \|\sum_1^n h(x_i)f_i\| < +\infty$; if $\{a_n\}$ is such that $\sup_n \|\sum_1^n a_i f_i\| < +\infty$ then there is a $h \in X^*$ with $h(x_i) = a_i$.*

A proof of (i) and (ii) may be found in [2] and of (iii) and (iv) in [19] (see also [20], cor. 3, p. 210, Lemma 2, p. 212 and prob. 27, p. 214).

3. The equivalences. Consider the following properties that a basis $\{x_i, f_i\}$ for a Banach space X may possess:

(B) $\lim_n \|V_n^*(f)\| = 0$ for each $f \in X^*$, where V^* is the adjoint of the operator V_n defined in 2.1 (ii);

(C) $\{f_i, Qx_i\}$ is a Schauder basis for X^*;

(D) $\{x_i, f_i\}$ is weakly uniform, i.e. $\sum_1^{\infty} f_i(x)x_i$ converges uniformly in the weak topology for x in the unit ball of X;

(E) A bounded sequence $\{y_n\}$ in X must converge weakly to zero whenever $\lim_n f_K(y_n) = 0$ for each $K \in \omega$;

(F) A bounded sequence $\{y_n\}$ in X must converge weakly to zero whenever $f_K(y_n) = 0$ for $K < n$;

(G) $\lim_n \operatorname{dist}(f, [f_i | i \leqq n]) = 0$ for each $f \in X^*$;

(H) $[f_m | m \neq n]$ is w^*-closed for each n;

(I) If $T = \{F \in X^{**} | \sum_1^{\infty} F(f_i)x_i \text{ converges}\}$ then $QX = T$;

(J) The mapping μ: X^{**} into E defined by $\mu(F) = \{F(f_i)\}$ is a linear homeomorphism of X^{**} onto E, where $E = \{\{a_i\} | \sup_n \|\sum_1^n a_i x_i\| < \infty\}$, with $\|\{a_i\}\| = \sup_n \|\sum_1^n a_i x_i\|$;

(K) $[f_n | n \in \omega] = X^*$; and,

(L) $\sup_n \|\sum_1^n a_i f_i\| < +\infty$ implies $\sum_1^{\infty} a_i f_i$ exists.

THEOREM 3.1. *Let* $\{x_i, f_i\}$ *be a basis for a Banach space* X. *Then the properties* (A) *through* (L) *are equivalent.*

Proof. (A)⇒(B). By 2.1 (ii) there is a K such that $\|V_n\| \leqq K$ for all $n \in \omega$. Thus, if $f \in X^*$,

$$\|V_n^* f\| = \sup\{\,|f(V_n(x))|\,|\,\|x\| \leqq 1\} \leqq \|f\|_n \cdot \sup\{V_n(x)\|\,|\,\|x\| \leqq 1\}$$
$$= \|f\|_n \|V_n\| \leqq K\|f\|_n.$$

By (A), $\lim_n \|f\|_n = 0$ and so (B) holds.

(B)⇒(C). Let I be the identity operator on X and I_0 the identity operator on X^*. Then $V_n = I - U_n$ for each n (see 2.1 (ii)), and if $f \in X^*$ it is clear that $U_n^*(f) = \sum_1^n f(x_i)f_i$. Thus,

$$\|V_n^*(f)\| = \|(I - U_n)^* f\| = \|(I_0 - U_n^*)f\| = \|f - \textstyle\sum_1^n f(x_i)f_i\|.$$

Thus if (B) holds, it follows that $f = \sum_1^\infty f(x_i)f_i = \sum_1^\infty Qx_i(f)f_i$. Since $Qx_i(f_i) = f_i(x_i) = \delta_{ji}$ it follows that the expansion is unique and so (C) holds.

(C)⇒(D). Let $U(0; g_1, \cdots, g_p; \epsilon)$ be a weak neighborhood of 0 in X. From (C) we obtain the existence of an N such that $n \geqq N$ implies $\|g_j - \sum_1^n g_j(x_i)f_i\| < \epsilon$, $j = 1, \cdots, p$ and so $|g_j(x) - \sum_1^n g_j(x_i)f_i(x)| < \epsilon$ if $\|x\| \leqq 1$, i.e. $\{x_i, f_i\}$ is weakly uniform.

(D)⇒(E). If $f \in X^*$ and $\epsilon > 0$ consider the weak neighborhood $U(0; f; \epsilon/2)$. Suppose that $\{y_n\}$ is a bounded sequence in X such that $\lim_n f_K(y_n) = 0$ for each $K \in \omega$. Without loss of generality we may suppose $\|y_n\| \leqq 1$. Since $\{x_i, f_i\}$ is weakly uniform there is an N such that $|f(y_n) - f(\sum_1^N f_i(y_n)x_i)| < \epsilon/2$ for all $n \in \omega$. Since $\lim_n f_K(y_n) = 0$ for each $K \in \omega$ there is a p such that $n \geqq p$ implies $|f(\sum_1^N f_i(y_n)x_i)| < \epsilon/2$. Therefore $|f(y_n)| < \epsilon$ if $n \geqq p$, i.e. $\{y_n\}$ converges weakly to zero.

(E)⇒(F). This implication is trivial.

(F)⇒(G). Suppose that (G) does not hold. Then there is an $f_0 \in X^*$ such that $\lim_n \operatorname{dist}(f_0, [f_i | i \leqq n]) \neq 0$. From the isometry $X^*/[f_i | i \leqq n] \equiv ([f_i | i \leqq n]\perp)^*$ we have $\operatorname{dist}(f_0, [f_i | i \leqq n]) = \sup\{|f_0(x)|\,|\,x \in [f_i | i \leqq n]\perp,\ \|x\| \leqq 1\}$. Since $f_i(x_j) = \delta_{ij}$ it follows readily that $[f_i | i \leqq n]\perp = [x_i | i \geqq n+1]$ and so $\operatorname{dist}(f_0, [f_i | i \leqq n]) = \sup\{|f_0(x)|\,|\,x \in [x_i | i \geqq n+1], \|x\| \leqq 1\}$. Thus, since $\lim_n \operatorname{dist}(f_0, [f_i i \leqq n]) \neq 0$, there is an $\epsilon > 0$, an increasing sequence of positive integers $\{n_N\}$, and $y_{n_N} \in [x_i | i \geqq n_N + 1]$, $\|y_{n_N}\| \leqq 1$ such that $|f_0(y_{n_N})| \geqq \epsilon$ and so $\{y_{n_N}\}$ does not converge weakly to zero. Clearly $f_K(y_{n_N}) = 0$ if $K < n_N$. Thus (F) does not hold.

(G)⇒(H). If (H) does not hold there is an n such that $M_n = [f_i | i \neq n]$ is not w^*-closed and so there is an $f \in X^* \backslash M_n$ such that $f(x) = 0$ for all $x \in (M_n)\perp$. Since $x_n \in (M_n)\perp$ we have $f(x_n) = 0$ and so $f \neq f_n$ since $f_n(x_n) = 1$. It follows that $f \notin [f_n | n \in \omega]$; for, if $f \in [f_n | n \in \omega]$ then by 2.1(i), $f = \sum_1^\infty f(x_i)f_i = \sum_{m \neq n} f(x_i)f_i$ since $f(x_n) = 0$. Thus $f \in M_n$, contradicting $f \in X^* \backslash M_n$.

On the other hand, $\operatorname{dist}(f, [f_n | n \in \omega]) \leqq \operatorname{dist}(f, [f_i | i \leqq n])$ and so if (G) holds we have $\operatorname{dist}(f, [f_n | n \in \omega]) = 0$, whence $f \in [f_n | n \in \omega]$. This contradiction shows that (G) implies (H).

(H)$\Rightarrow$(I). If $M_n = [f_i | i \neq n]$ is w^*-closed for each n it follows that $[f_i | i \in \omega]$ is w^*-closed. But $\{f_i\}$ is total over X and so the span of $\{f_i\}$ is w^*-dense in X^*; thus $[f_i | i \in \omega]$ is all of X^*. Thus by 2.1(i) for each $f \in X^*$, $f = \sum_1^\infty f(x_i) f_i$ (convergence in the norm topology). If $F \in X^{**}$ is such that $\sum_1^\infty F(f_i) x_i = x_F$ exists then for each $f \in X^*$ we have

$$(F - Qx_F)f = F(f) - f(x_F) = F(f) - f\left(\sum_1^\infty F(f_i)x_i\right)$$

$$= F(f) - F\left(\sum_1^\infty f(x_i)f_i\right) = 0.$$

Therefore $F = Qx_F$ and so $T \subset QX$. On the other hand, since $\{x_i, f_i\}$ is a basis for X we have $QX \subset T$ and (I) holds.

(I)$\Rightarrow$(J). Suppose $F \in X^{**}$ and $\mu(F) = 0$. Then $F(f_i) = 0$ for each i and so $\sum_1^\infty F(f_i) x_i = 0$. From (I) it follows that $F = Qx_F$ for some $x_F \in X$. But $0 = F(f_i) = Qx_F(f_i) = f_i(x_F)$ for each i and so $x_F = 0$. Thus $F = 0$ and so μ is one-to-one. That μ is onto follows from 2.1(iii). Also, for $F \in X^{**}$ we have

$$\|\mu(F)\| = \sup_n \left\| \sum_1^n F(f_i) x_i \right\| = \sup_n \sup_{\|f\| \leqq 1} \left| \sum_1^n F(f_i) f(x_i) \right|$$

$$\leqq \|F\| \sup_n \sup_{\|f\| \leqq 1} \left\| \sum_1^n f(x_i) f_i \right\| = \|F\| \sup_n \|U_n^*\| \leqq K\|F\| \quad \text{by 2.1(ii)},$$

since $\|U_n^*\| = \|U_n\|$ for each n. Thus μ is continuous and hence, by the open mapping theorem, a linear homeomorphism onto E.

(J)$\Rightarrow$(K). If $[f_n | n \in \omega] \neq X^*$ then, by the Hahn-Banach theorem, there is an $F \in X^{**}$ such that $F(f_i) = 0$ for each i and $F \neq 0$. Thus μ is not one-to-one and so (J) does not hold.

(K)$\Rightarrow$(L). If $\sup_n \|\sum_1^n a_i f_i\| < +\infty$ then by 2.1(iv) there is an $h \in X^*$ such that $a_i = h(x_i)$ for all i. By 2.1(i) and (K) we have $h = \sum_1^\infty h(x_i) f_i = \sum_1^\infty a_i x_i$ and so (L) holds.

(L)$\Rightarrow$(A). If $h \in X^*$ then by 2.1(iv) $\sup_n \|\sum_1^n h(x_K) f_K\| < +\infty$ and so by (L), $\sum_1^\infty h(x_K) f_K$ exists and clearly must be equal to h. For $\epsilon > 0$ choose $y_n = \sum_n^\infty a_i x_i \in [x_i | i \geqq n]$ such that $\|y_n\| = 1$ and $|h(y_n)| > \|h\|_n - \epsilon/2$.

Then $|h(y_n)| = |\sum_1^\infty h(x_K) f_K(\sum_n^\infty a_i x_i)| = |\sum_n^\infty h(x_K) f_K(\sum_n^\infty a_i x_i)| \leqq (\|\sum_n^\infty h(x_K) f_K)\|$ since $\|y_n\| = 1$. If N is chosen such that $n \geqq N$ implies $\|\sum_n^\infty h(x_K) f_K\| < \epsilon/2$ then $\|h\|_n < |h(y_n)| + \epsilon/2 \leqq \|\sum_n^\infty h(x_K) f_K\| + \epsilon/2 < \epsilon$. Thus (A) holds, completing the proof of the theorem.

4. Remarks and unsolved problems. In [18] Singer proved the following theorem:

THEOREM 4.1. *If $\{x_i, f_i\}$ is a shrinking basis for a Banach space X then X^* is linearly homeomorphic by the mapping*

$$\nu: f \to \{f(x_i)\} \quad to \quad F = \left\{ \{a_n\} \;\middle|\; \sup_n \left\| \sum_1^n a_i f_i \right\| < +\infty \right\},$$

where the norm on F is defined by $\|\{a_n\}\| = \sup_n \|\sum_1^n a_i f_i\|$.

We observe here that the hypothesis that the basis be shrinking in 4.1 is extraneous, i.e.

If $\{x_i, f_i\}$ is a basis for a Banach space X then ν is a linear homeomorphism from X^* onto F where ν and F are as defined in 4.1.

To see that ν is one-to-one observe that $\nu(f) = 0$ implies $f(x_i) = 0$ for each i and so, since $\{x_i, f_i\}$ is a basis for X, $f = 0$. That ν is onto follows from 2.1(iv).

If $f \in X^*$ then

$$\|\nu(f)\| = \sup_n \left\| \sum_1^n f(x_i) f_i \right\| = \sup_n \sup_{\|x\| \leqq 1} \left| \sum_1^n f(x_i) f_i(x) \right|$$
$$\leqq \|f\| \sup_n \|U_n\| \leqq K\|f\|$$

by 2.1(ii). Thus ν is continuous, and so, by the open mapping theorem, a linear homeomorphism of X^* onto F.

Of course in the preceding proof we have tacitly assumed that F is a Banach space. The only property not immediately obvious is that F is complete. We sketch a proof of the completeness.

Let $\{a^{(n)}\}$ be a Cauchy sequence in F, $a^{(n)} = \{a_i^{(n)}\}$, then from the definition of the norm in F it follows that, for each i, $\{a_i^{(n)}\}$ is a Cauchy sequence of scalars. Let $a_i = \lim_{n\to\infty} a_i^{(n)}$. Since $\{a^{(n)}\}$ is Cauchy in F there is a positive real number M such that $\|a^{(n)}\| \leqq M$ for every n. Let K be a fixed positive integer. There is an N_K such that $n > N_K$ implies $\|\sum_1^K (a_i - a_i^{(n)}) f_i\| < 1$ and so

$$\left\| \sum_1^K a_i f_i \right\| \leqq \left\| \sum_1^K (a_i - a_i^{(n)}) f_i \right\| + \|a^{(n)}\| \leqq M + 1 \quad \text{if } n > N_K.$$

Thus $\sup_K \|\sum_1^K a_i f_i\| \leqq M + 1$ whence $a = \{a_i\} \in F$. Clearly $\lim_{n\to\infty} a^{(n)} = a$, and F is complete.

Finally we list two unsolved problems related to the material in this paper.

P1. If X^* is separable and X has a basis does X have a shrinking basis?

P2. Does there exist a nonreflexive Banach space X with a basis such that every basis in X is shrinking?

In connection with P2 it is known that a Banach space X is reflexive if and only if every basic sequence in X is shrinking (see [8] and [14]).

References

1. Y. Cuttle, On quasi-reflexive Banach spaces, Proc. Amer. Math. Soc., 12 (1961) 936–940.
2. M. M. Day, Normed linear spaces, Springer-Verlag, Berlin, 1958, p. 67–77.
3. B. R. Gelbaum, Expansions in Banach spaces, Duke Math. J., 17(1950) 187–196.
4. R. C. James, Bases and reflexivity of Banach spaces. Ann. of Math., 52(1950) 518–527.
5. ———, A non-reflexive Banach space isometric with its second conjugate space, Proc. Nat. Acad. Sci. USA, 37(1951) 174–177.
6. ———, Projections in the space (m), Proc. Amer. Math. Soc., 6 (1955) 899–902.
7. ———, Separable conjugate spaces, Pacific J. Math., 10(1960) 563–571.
8. A. Pelczynski, A note on the paper of I. Singer "Basic sequences and reflexivity of Banach spaces", Studia Math., 21(1962) 371–374.
9. J. R. Retherford, w* and bw*-bases in Banach spaces, 25 (1964) 65–71.
10. ———, Bases, basic sequences and reflexivity of linear topological spaces, Math. Ann., 64 (1966) 280–285.
11. ———, and C. W. McArthur, Uniform and equicontinuous Schauder bases of subspaces, Canadian J. Math., 7 (1965) 207–212.
12. W. H. Ruckle, The infinite sum of closed subspaces of an F-space, Duke Math. J., 31(1964) 543–554.
13. B. L. Sanders, Decompositions and reflexivity, Proc. Amer. Math Soc., 6 (1965) 204–208.
14. I. Singer, Basic sequences and reflexivity of Banach spaces, Studia Math., 21(1962) 351–369.
15. ———, Bases and quasi-reflexivity of Banach spaces, Math. Ann., 153(1964) 199–209.
16. ———, On bases in quasi-reflexive Banach spaces, Rev. Math. Pures Appl., 8(1963) 309–311.
17. ———, On Banach spaces reflexive with respect to a linear subspace of their conjugate space, II, Math. Ann., 145 (1961/62) 64–76.
18. ———, On Banach spaces reflexive with respect to a linear subspace of their conjugate space, III, Rev. Math. Pures Appl., 7(1962) 135–142.
19. A. Wilansky, The basis in Banach space, Duke Math. J., 18(1951) 795–798.
20. ———, Functional Analysis, Blaisdell, New York, 1964, p. 210–214.

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457

ON THE MOESSNER THEOREM ON INTEGRAL POWERS

C. T. LONG, Washington State University

1. Introduction. In a paper [1, p. 29] in 1951, Alfred Moessner conjectured that the natural kth powers can be generated in the following interesting way: From the series of natural numbers strike out every kth number. From the resulting series form the series of partial sums and from this series strike out every $(k-1)$-st number. Again form the series of partial sums and now delete every $(k-2)$-nd number. Repeat the process $k-1$ times, striking out every second number the last time and forming the final series of partial sums. This final

series of partial sums is the series of natural kth powers. For example, for $k=4$ we obtain

$$\begin{array}{cccccccccccccc}
1 & 2 & 3 & \cancel{4} & 5 & 6 & 7 & \cancel{8} & 9 & 10 & 11 & \cancel{12} & \cdots \\
1 & 3 & \cancel{6} & & 11 & 17 & \cancel{24} & & 33 & 43 & \cancel{54} & & \cdots \\
1 & \cancel{4} & & & 15 & \cancel{32} & & & 65 & \cancel{108} & & & \cdots \\
1 & & & & 16 & & & & 81 & & & & \cdots
\end{array}$$

and we note that $1=1^4$, $16=2^4$, and $81=3^4$. That Moessner's conjecture is true was first proved by Oskar Perron [2, pp. 31–34] using the method of mathematical induction. I. Paasche [3, pp. 1–5] derived Moessner's theorem as a consequence of general theorems on the theory of generating functions and H. Salié [4, pp. 7–11] used mathematical induction to prove a theorem concerning the application of the Moessner process to an arbitrary sequence.

In this paper we use mathematical induction and an interesting generalization of Pascal's triangle to prove that if Moessner's process is applied to the arithmetic progression

$$a,\ a+d,\ a+2d,\ a+3d,\ \cdots,$$

the final series of partial sums is given by

$$a\cdot 1^{k-1},\ (a+d)\cdot 2^{k-1},\ (a+2d)\cdot 3^{k-1},\ \cdots.$$

The same method could be used to prove Salié's theorem from which, in turn, the present result can be derived. Of course, if $a=d=1$, the present result reduces to Moessner's theorem.

2. Generalization of Pascal's triangle. For $n\geqq 0$, $0\leqq n\leqq r$, let (n, r) denote the rth element in the nth row of the array

$$\begin{array}{ccccccccccc}
 & & & & & b_0 & & & & & \\
 & & & & a_1 & & b_1 & & & & \\
 & & & a_2 & & a_1+b_1 & & b_2 & & & \\
 & & a_3 & & a_2+a_1+b_1 & & a_1+b_1+b_2 & & b_3 & & \\
 & a_4 & & a_3+a_2+a_1+b_1 & & a_2+2a_1+2b_1+b_2 & & a_1+b_1+b_2+b_3 & & b_4 & \\
a_5 & & a_4+a_3+a_2+a_1+b_1 & & a_3+2a_2+3a_1+3b_1+b_2 & & a_2+3a_1+3b_1+2b_2+b_3 & & a_1+b_1+b_2+b_3+b_4 & & b_5 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}$$

defined recursively by

$$\begin{aligned}
(n, 0) &= a_n, && n\geqq 1,\\
(n, n) &= b_n, && n\geqq 0, \qquad (1)\\
(n, r)+(n, r+1) &= (n+1, r+1), && 0\leqq r\leqq n-1,\ n\geqq 1.
\end{aligned}$$

Extending the array, it is not difficult to guess that

$$(2)\qquad (n, r)=\sum_{i=1}^{n-r}\binom{n-1-i}{r-1}a_i+\sum_{j=1}^{r}\binom{n-1-j}{r-j}b_j \qquad n\geqq 2,\ 1\leqq r\leqq n-1.$$

This is clearly true for $n=2$. Suppose that it is also true for $n=m$. Then, by (1),

$$\begin{aligned}
(m+1,1) &= (m,0)+(m,1)\\
&= a_m+\sum_{i=1}^{m-1}\binom{m-1-i}{0}a_i+\sum_{j=1}^{1}\binom{m-1-j}{1-j}b_j\\
&= \sum_{i=1}^{m}\binom{m-i}{0}a_i+\sum_{j=1}^{1}\binom{m-j}{1-j}b_j,\\
(m+1,m) &= (m,m-1)+(m,m)\\
&= \sum_{i=1}^{1}\binom{m-1-i}{m-2}a_i+\sum_{j=1}^{m-1}\binom{m-1-j}{m-1-j}b_j+b_m\\
&= \sum_{i=1}^{1}\binom{m-i}{m-1}a_i+\sum_{j=1}^{m}\binom{m-j}{m-j}b_j,
\end{aligned}$$

and, for $2\leqq r\leqq m-1$,

$$\begin{aligned}
(m+1,r) &= (m,r-1)+(m,r)\\
&= \sum_{i=1}^{m-r+1}\binom{m-1-i}{r-2}a_i+\sum_{j=1}^{r-1}\binom{m-1-j}{r-1-j}b_j\\
&\quad+\sum_{i=1}^{m-r}\binom{m-1-i}{r-1}a_i+\sum_{j=1}^{r}\binom{m-1-j}{r-j}b_j\\
&= a_{m-r+1}+\sum_{i=1}^{m-r}\left\{\binom{m-1-i}{r-2}+\binom{m-1-i}{r-1}\right\}a_i\\
&\quad+b_r+\sum_{j=1}^{r-1}\left\{\binom{m-1-j}{r-1-j}+\binom{m-1-j}{r-j}\right\}b_j\\
&= \sum_{i=1}^{m-r+1}\binom{m-i}{r-1}a_i+\sum_{j=1}^{r}\binom{m-j}{r-j}b_j.
\end{aligned}$$

Thus, (2) holds for $n=m+1$ if it holds for $n=m$ and so, by mathematical induction, it holds for all $n\geqq 2$ as claimed.

3. Proof of the main result. To see how the discussion of the preceding paragraph applies to the problem at hand, consider the following array for the case $k=4$:

$$(3)\qquad \begin{array}{llllllll}
a & a+d & a+2d & \cancel{a+3d} & a+4d & a+5d & a+6d & \cancel{a+7d}\cdots\\
a & 2a+d & \cancel{3a+3d} & & 4a+7d & 5a+12d & \cancel{6a+18d} & \cdots\\
a & \cancel{3a+d} & & & 7a+8d & \cancel{12a+20d} & & \cdots\\
a & & & & 8a+8d & & &
\end{array}$$

It is clear that this is simply a sequence of triangular arrays each constructed

like the generalized Pascal triangle above. Alternatively, we may consider the array:

$$
(4)\qquad
\begin{array}{ccccccccccc}
d & d & d & d & \cancel{d} & d & d & d & d & \cancel{d} & \cdots \\
a & a+d & a+2d & \cancel{a+3d} & & a+4d & a+5d & a+6d & \cancel{a+7d} & & \cdots \\
a & 2a+d & \cancel{3a+3d} & & & 4a+7d & 5a+12d & \cancel{6a+18d} & & & \\
a & \cancel{3a+d} & & & & 7a+8d & \cancel{12a+20d} & & & & \\
a & & & & & 8a+8d & & & & &
\end{array}
$$

obtained from (3) by the obvious addition of the row of d's. Again the triangular arrays involved are generalized Pascal triangles, but here the elements in the first column of any particular array are the partial sums of the elements in the last diagonal row of the preceding array. Thus, in the second triangular array in (4)

$$
\begin{aligned}
d &= d, \\
a + 4d &= d + (a + 3d), \\
4a + 7d &= d + (a + 3d) + (3a + 3d),
\end{aligned}
$$

and so on.

In general, we consider an array like (4), constructed by the Moessner process by initially deleting every $(k+1)$-st d with $k \geqq 2$. For $t \geqq 1$, $n \geqq 0$, $0 \leqq r \leqq n$, let $(n, r)_t$ denote the rth element (reading from upper right to lower left) in the nth diagonal of the tth triangle in the resulting array. In (4), for example, we have that $(3, 0)_1 = d$, $(3, 1)_1 = a + 2d$, $(4, 2)_2 = 6a + 18d$, and so on. In these terms, then, our claim is that

$$
(5) \qquad (k, k)_t = t^{k-1} \cdot \{a + (t - 1)d\} \qquad k \geqq 1,\ t \geqq 1.
$$

In fact, more generally, we claim that

$$
(6) \qquad (k, r)_t = (a - d)\binom{k-1}{r-1} t^{r-1} + d\binom{k}{r} t^r, \qquad k \geqq 1,\ t \geqq 1,\ 0 \leqq r \leqq k,
$$

where we agree to set

$$
\binom{k-1}{-1} = 0.
$$

For $t = 1$ we obtain (6) from the results of Section 2 with $n = k$, $a_i = d$ for $1 \leqq i \leqq k$, and $b_j = a$ for $1 \leqq j \leqq k$. Thus, $(k, 0)_1 = d$, $(k, k)_1 = a$, and, for $1 \leqq r \leqq k - 1$, we see from (2) that

$$
(k, r)_1 = \sum_{i=1}^{k-r} \binom{k-1-i}{r-1} d + \sum_{j=1}^{r} \binom{k-1-j}{r-j} a
$$

$$= \binom{k-1}{r} d + \binom{k-1}{r-1} a$$

$$= \left\{ \binom{k-1}{r} + \binom{k-1}{r-1} \right\} \cdot d + \binom{k-1}{r-1} (a-d)$$

$$= \binom{k}{r} d + \binom{k-1}{r-1} (a-d).$$

Thus, (6) holds for $t=1$.

Suppose that (6) also holds for $t=m$; i.e., suppose that

$$(7) \qquad (k, r)_m = \binom{k-1}{r-1} m^{r-1} \cdot (a-d) + \binom{k}{r} m^r \cdot d \qquad 0 \leqq r \leqq k.$$

Then, since the partial sums of the diagonal elements of the mth triangular array form the elements in the first column of the $(m+1)$-st array, in determining the values of $(k, r)_{m+1}$ we use the results of Section 2 with $a_i = d$ for $1 \leqq i \leqq k$ and, by (7),

$$(8) \qquad \begin{aligned} b_j &= \sum_{s=0}^{j} (k, s)_m \\ &= \sum_{s=0}^{j} \left\{ (a-d) \binom{k-1}{s-1} m^{s-1} + d \binom{k}{s} m^s \right\} \end{aligned}$$

for $1 \leqq j \leqq k$. Thus, by (1), $(k, 0)_{m+1} = a_k = d$ and

$$\begin{aligned} (k, k)_{m+1} &= b_k \\ &= \sum_{s=0}^{k} \left\{ (a-d) \binom{k-1}{s-1} m^{s-1} + d \binom{k}{s} m^s \right\} \\ &= (a-d)(m+1)^{k-1} + d(m+1)^k. \end{aligned}$$

Moreover, since $d = a_i$ and we are taking

$$\binom{k-1}{-1}$$

to be zero, we have from (2) and (8) that

$$\begin{aligned} (k, r)_{m+1} = &\sum_{i=1}^{k-r} \binom{k-1-i}{r-1} d, \\ &+ \sum_{j=1}^{r} \binom{k-1-j}{r-j} \sum_{s=0}^{j} \left\{ (a-d) \binom{k-1}{s-1} m^{s-1} + d \binom{k}{s} m^s \right\} \end{aligned}$$

$$
\begin{aligned}
&= d\binom{k-1}{r} + (a-d)\sum_{j=1}^{r}\sum_{s=1}^{j}\binom{k-1-j}{r-j}\binom{k-1}{s-1}m^{s-1} \\
&\quad + d\sum_{j=1}^{r}\sum_{s=0}^{j}\binom{k-1-j}{r-j}\binom{k}{s}m^{s} \\
&= (a-d)\sum_{s=1}^{r}\sum_{j=s}^{r}\binom{k-1-j}{r-j}\binom{k-1}{s-1}m^{s-1} \\
&\quad + d\sum_{s=0}^{r}\sum_{j=s}^{r}\binom{k-1-j}{r-j}\binom{k}{s}m^{s} \\
&= (a-d)\sum_{s=1}^{r}\binom{k-1}{s-1}\binom{k-s}{r-s}m^{s-1} + d\sum_{s=0}^{r}\binom{k}{s}\binom{k-s}{r-s}m^{s} \\
&= (a-d)\binom{k-1}{r-1}\sum_{s=1}^{r}\binom{r-1}{s-1}m^{s-1} + d\binom{k}{r}\sum_{s=0}^{r}\binom{r}{s}m^{s} \\
&= (a-d)\binom{k-1}{r-1}(m+1)^{r-1} + d\binom{k}{r}(m+1)^{r}
\end{aligned}
$$

for $1 \leqq r \leqq k-1$.

Thus, (6) is true for $t=m+1$ if it is true for $t=m$ and the result follows by mathematical induction.

References

1. A. Moessner, Eine Bemerkung über die Potenzen der natürlichen Zahlen, Bayer. Akad. Wiss. Math.-Nat. Kl. S.-B., (1951) 29.
2. O. Perron, Beweis des Moessnerschen Satzes, *ibid.*, (1951) 31–34.
3. I. Paasche, Ein neuer Beweis des Moessnerschen Satzes, *ibid.*, (1952) 1–5.
4. H. Salié, Bemerkung zu einem Satz von A. Moessner, *ibid.*, (1952) 7–11.

SOME REMARKS ON THE PARTITION FUNCTION

M. V. Subbarao, University of Alberta, Edmonton

1. Preliminaries. Kolberg [2] was probably the first to prove that the partition function $p(n)$ (which denotes the number of unrestricted partitions of the positive integer n) takes both odd and even values, each of them infinitely often. This result is also contained in the Advanced problem No. 4944 (1961, 76) proposed by M. Newman and solved by J. H. van Lint and the proposer (1962, 69, p. 175) in this Monthly. As a wide generalization of this result M. Newman [3] conjectured that for all integers $m \geqq 2$ each of the congruences

$$p(n) \equiv r \pmod{m}, \qquad 0 \leqq r \leqq m-1,$$

his infinitely many solutions in positive integers n. He proved this conjecture for $m=2, 5, 13, 65$ ([3], [4]).

In this note we prove that $p(2n+1)$ takes even and odd values, each of them infinitely often. It should be noted that our result implies Kolberg's result and yet is not covered by Newman's conjecture. We also prove some other congruences and end up with some conjectures.

2. A recurrence formula with a congruence. In the sequel we write

$$\phi(x) = \prod_{n=1}^{\infty} (1 - x^n), \qquad |x| < 1;$$

$\sigma(n)$ and $t(n)$ for the sum and number of divisors of n respectively, and define $\sigma(0)=0$; $t(0)=p(0)=1$, $p(n)=0$ for $n<0$. Also we write θ for the operator xd/dx and recall that

$$1/\phi(x) = \sum p(n)x^n, \tag{2.1}$$

$$\theta \log(1/\phi(x)) = \sum \sigma(n)x^n, \tag{2.2}$$

where, unless otherwise specified, all summations are from $n=0$ to $n=\infty$.

In addition to the classical result due to Euler on the power series expansion for $\phi(x)$ we use the following well-known and easily proved result:

LEMMA 1. *For all positive integers m and all divisors $d>2$ of 24,*

$$\sigma(dm - 1) \equiv 0 \pmod{d}. \tag{2.3}$$

This holds for $d=2$ if and only if $dm-1$ is not a square integer. (Of course (2.3) holds trivially for $d=1$.) A simple proof of this result, given in [**5**], is based on the fact that if $d>2$, the integer $N=dm-1$ is nonsquare, and for any $t|N$, $t+(N/t)\equiv 0 \pmod{d}$. This leads to

LEMMA 2. *Let $d\geqq 2$ be a divisor of 24. For all nonsquare N for which*

$$N \equiv -1 \pmod{d} \tag{2.4}$$

we have

$$\begin{aligned} p(N) \equiv \sum_{k>0} (-1)^{k-1} \Bigg\{ &\left(1 + \frac{k(3k-1)}{2}\right) p(N - k(3k-1)/2) \\ &+ \left(1 + \frac{k(3k+1)}{2}\right) p(N - k(3k+1)/2) \Bigg\} \pmod{d}. \end{aligned}$$

REMARK. If $d>2$, any N satisfying (2.4) is a nonsquare integer.

Proof. From (2.1) and (2.2) we have

$$\phi(x) \sum np(n)x^n = \sum \sigma(n)x^n.$$

Equating the coefficients of x^N on both sides and using Lemma 1 we have the required result.

COROLLARY. *For N, d as given above,*

$$(2.5) \quad p(N-1)+2p(N-2)-5p(N-5)-7p(N-7)+\cdots \equiv 0 \pmod{d},$$

the general term on the left being $(-1)^{k-1}k(3k\pm 1)/2\ p(N-k(3k\pm 1)/2)$.

This follows on combining the result of Lemma 2 with Euler's power series expansion for $\phi(x)$. In particular, setting $d=3$ and observing that $k(3k-1)/2 \equiv k \pmod 3$ and $k(3k+1)/2 \equiv -k \pmod 3$ we have for *all* $N\equiv 2 \pmod 3$,

$$\sum_{k\geqq 1, 3\nmid k} (-1)^{[k/3]}\{p(N-k(3k-1)/2)-p(N-k(3k+1)/2)\} \equiv 0 \pmod{3}.$$

3. Parity of $p(2n+1)$.

THEOREM 1. *The congruences $p(2n+1)\equiv 0 \pmod 2$ and $p(2n+1)\equiv 1 \pmod 2$ have each an infinite number of solutions in n.*

Proof. We apply an argument similar to Kolberg's to the relation

$$(3.1) \quad p(N)+p(N-2)+p(N-12)+p(N-22)+\cdots \equiv 0 \pmod{2},$$

which is valid for any positive odd integer N which is not a perfect square, where in the left member of (3.1) the general term is $p(N-k(3k\pm 1)/2)$ provided $k(3k\pm 1)/2$ is even. This relation is a consequence of Lemma 2.

Assuming, then, that the theorem is false, let m be the largest odd integer for which $p(m)$ is even (odd). Setting $r=32m+8$ and $N=m+r(3r-1)/2$ it is easily verified that $N\equiv 2 \pmod 3$ and hence not a square integer. Further, since N is odd, the congruence (3.1) applies and the last term in the left member of the congruence is $p(m)$. Setting $a_k=k(3k-1)/2$, $b_k=k(3k+1)/2$, the terms of the sequence $a_1, b_1, a_2, b_2, \cdots$ taken modulo 2 recur, the recurrence cycle being 1, 0, 1, 1, 0, 1, 0, 0. Recalling the values of r and N in terms of m it is seen that in the left member of (3.1) the number of terms is even, being $32m+8$; of these the only even (odd) term is $p(m)$. Hence for this choice of N the congruence (3.1) cannot hold, thus proving the theorem.

4. The function $s_r(n)$. Let $a_1, a_2, \cdots, a_{p(n)}$ denote the number of parts (summands) in the $p(n)$ partitions of n arranged in some order. We define $s_r(n)=a_1^r+\cdots+a_{p(n)}^r$ so that $s_0(n)=p(n)$. We shall write $s(n)$ for $s_1(n)$. Let

$$\phi(x,b)=(1-bx)(1-bx^2)\cdots$$

so that $1/\phi(x,b)=1+B_1x+B_2x^2+\cdots$, where $B_n=b^{a_1}+\cdots+b^{a_{p(n)}}$. It is easily seen, using differentiation with respect to b a couple of times and setting $b=1$ (and defining $s_r(n)=0$ for $n\leqq 0$) that

$$(4.1) \quad \begin{aligned} \phi(x)\sum s(n)x^n &= \sum t(n)x^n, \\ \phi(x)\sum s_2(n)x^n &= \left(\sum t(n)x^n\right)^2+\sum \sigma(n)x^n. \end{aligned}$$

Using the Euler expansion for $\phi(x)$, the congruence properties for $\sigma(n)$ in Lemma

1 and the fact that $t(n) \equiv 0 \pmod 2$ unless n is a square we can prove not only recurrence relations for $s(n)$ and $s_2(n)$ as in Lemma 2, but also:

THEOREM 2. *The functions $s(n)$ and $s_2(n)$ take infinitely often even as well as odd values.*

The proof of this is analogous to that of Theorem 1. The function $s(n)$ occurs in a recent paper of Fine [1], where (4.1) is obtained by a different method. A generating series for $s_r(n)$ for general values of r seems to be fairly complicated.

5. Concluding remarks. The author has been unable to prove the analogue of Theorem 1 for $p(2n)$—namely that, for infinitely many n, $p(2n)$ is odd (even), but believes it to be true. It might be conjectured that $s_r(n)$ takes both even and odd values, each infinitely often, for all integers $r \geqq 0$, and probably more, that for all integers $k \geqq 1$ and all a, $0 \leqq a < k$, each of the congruences $s_r(nk+a) \equiv 0 \pmod k$, $s_r(nk+1) \equiv 1 \pmod k$ hold for an infinite number of integers n and all integers $r \geqq 0$.

In a recent private communication to the author, Professor O. Kolberg proved that $p(2n)$ takes even as well as odd values infinitely often. Following Kolberg's method, the author has now proved similar results concerning $p(4n+r)$ $(r=0, 1, 2, 3)$. These, as well as further extensions, will appear elsewhere.

References

1. N. J. Fine, Sums over partitions, Report of the Institute in the Theory of Numbers. University of Colorado, Boulder, (1959) pp. 86–94.

2. Oddmund Kolberg, Note on the parity of the partition function, Math. Scand., 7 (1959) 377–378.

3. Morris Newman, Periodicity modulo m and divisibility properties of the partition function, Trans. Amer. Math. Soc., 97 (1960) 225–236.

4.———, Congruences for the partition function to composite moduli, Illinois J. of Math., 6(1962) 59–63.

5. M. V. Subbarao, Congruence properties of $\sigma(n)$, Math. Student, (1950) 17–18.

SOME THEOREMS ON EXPANSIVE HOMEOMORPHISMS

RICHARD WILLIAMS, Southern Methodist University

In this note, a new theorem concerning expansive homeomorphisms will be proved, and an elementary proof of a known theorem appears as a corollary to the new theorem mentioned above.

The following theorem was first proved by Jakobsen and Utz in [4]. We now give a new proof.

THEOREM 1. *If D is the closed unit disk, then there does not exist an expansive homeomorphism on D.*

Proof. Let f be an expansive homeomorphism on D with expansive constant δ. Let C be the boundary of D. Since f maps C onto itself, f is expansive on C.

Also, since f is expansive if and only if f^n is expansive for $n \neq 0$ if f has a periodic point, we may assume that the point is fixed, ([5], Theorem 2.2).

Suppose x is fixed under f. Then f is expansive on $C-\{x\}$, and there is a homeomorphism g of $C-\{x\}$ onto (0, 1) such that g^{-1} is uniformly continuous. Hence gfg^{-1} is expansive on (0, 1), a contradiction, ([1], Theorems 1, 3 and 4). Therefore f has no periodic points.

There exists η such that $0<\eta<\min(\delta, 1)$ and such that $d(x, y)<\eta$ implies $d(f(x), f(y))<1$. Also, there exist a pair of positively asymptotic points, ([2], Theorem 2). Hence there exist a and b such that $n \geqq 0$ implies $d(f^n(a), f^n(b))<\eta$. Let A be the arc from a to b of diameter less than η. Then it follows by induction that $f^k(A)$ is the arc from $f^k(a)$ to $f^k(b)$ of diameter less than η for $k=0, 1, \cdots$. Thus each two points of A are positively asymptotic, since $x, y \in A$ implies $d(f^k(x), f^k(y))<\eta<\delta$ for $k=0, 1, \cdots$, (see [2], Lemma 1).

We will now show that $m \neq n$ implies $f^m(A) \cap f^n(A)=\phi$. If this is not true, there exist $x, y \in A$ such that $f^m(x)=f^n(y)$, i.e., $y=f^{m-n}(x)$. Since $x, y \in A$, $d(f^k(x), f^k(y))=d(f^k(x), f^{m-n+k}(x)) \rightarrow 0$ as $k \rightarrow \infty$. There exists a subsequence $\{k_i\}$ and a point $p \in C$ such that $f^{k_i}(x) \rightarrow p$. Then $f^{m-n+k_i}(x) \rightarrow f^{m-n}(p)$, so $f^{m-n}(p)=p$, a contradiction to the assumption that f has no periodic points.

Since $k \geqq 0$ implies $\operatorname{diam}(f^k(A))<\delta$, for each integer $N<0$ there exist a pair of points in A so close that they never get δ apart under f^n for $n \geqq N$. Thus there is a sequence $0>n_1>n_2> \cdots$ such that $\operatorname{diam}(f^{n_i}(A))>\delta$ for $i=1, 2, \cdots$. But $i \neq j$ implies $f^{n_i}(A) \cap f^{n_j}(A)=\phi$. Thus we have an infinite number of disjoint arcs, each of diameter greater than $\delta>0$, a contradiction.

Hence f is not expansive on C, and is therefore not expansive on D.

In Jakobsen and Utz' proof of the preceding theorem, they stated that the set of points doubly asymptotic to a fixed point is at most countable. A stronger statement can be made.

Henceforth, we assume that f is an expansive homeomorphism on X with expansive constant δ.

THEOREM 2. *If X is compact, then the set of points doubly asymptotic to a given point is at most countable.*

Proof. Let x be given. Then y is doubly asymptotic to x if and only if there exists a positive integer N such that $|n| \geqq N$ implies $d(f^n(x), f^n(y))<\delta/2$. Let $A_N=\{y: |n| \geqq N \text{ implies } d(f^n(x), f^n(y))<\delta/2\}$. Then the set of points doubly asymptotic to x is $\bigcup_{N=1}^{\infty} A_N$. Suppose some A_N, say A_M, is infinite. There exists $\eta>0$ such that $d(y, z)<\eta$ implies $d(f^n(y), f^n(z))<\delta$ for $|n|<M$. Since A_M is infinite, there exist $y, z \in A_M$ such that $d(y, z)<\eta$. Then $d(f^n(y), f^n(z))<\delta$ for each n, a contradiction. Thus each A_N is finite, and $\bigcup_{N=1}^{\infty} A_N$ is at most countable.

The method used in the proof of Theorem 2 can also be used to prove the following theorem:

THEOREM 3. *If X is complete and self-dense, then for each $x \in X$ and for each neighborhood N of x, there exists $y \in N$ such that x and y are not doubly asymptotic.*

Proof. Suppose that each $y \in N$ is doubly asymptotic to x. There exists a compact perfect set $A \subseteq N$ ([**3**], p. 103, exercise 2–46). Using the notation of Theorem 2, $A \subseteq \bigcup_{N=1}^{\infty} A_N$. Since A is uncountable, A intersects some A_N in an uncountable number of points. But since A is compact, the same reasoning used in the proof of Theorem 2 shows that each $A \cap A_N$ must be finite. We thus have a contradiction, and the theorem is proved.

The following theorem of Bryant ([**2**], Theorem 3) is an obvious corollary to both Theorems 2 and 3:

COROLLARY. *If X is compact and self-dense, then for each $x \in X$ and for each neighborhood N of x, there exists $y \in N$ such that x and y are not doubly asymptotic.*

References

1. B. F. Bryant, Expansive self-homeomorphisms of a compact metric space, this MONTHLY, 69 (1962) 386–391.
2. ———, On expansive homeomorphisms, Pacific J. Math., 10 (1960) 1163–1167.
3. John G. Hocking and Gail S. Young, Topology, Addison-Wesley, Reading, Mass., 1961.
4. J. F. Jakobsen and W. R. Utz, The nonexistence of expansive homeomorphisms on a closed 2-cell, Pacific J. Math., 10 (1960) 1319–1321.
5. W. R. Utz, Unstable homeomorphisms, Proc. Amer. Math. Soc., 1 (1950) 769–774.

ON A THEOREM OF POLYA

R. KELLNER, Case Institute of Technology

Introduction. In [**1**] Polya considers the two classical boundary-value problems

$$(1) \qquad \Delta u + \lambda u = 0 \quad \text{in } D, \qquad u = 0 \text{ on } C;$$

$$(2) \qquad \Delta u + \mu u = 0 \quad \text{in } D,\ \partial u/\partial n = 0 \text{ on } C;$$

D is a simply connected bounded plane domain, C its boundary curve, Δ the Laplace operator, and n the normal to C. Corresponding to the problems (1) and (2), respectively, are the two sequences of eigenvalues

$$0 < \lambda_1 < \lambda_2 \leqq \lambda_3 \leqq \lambda_4 \leqq \cdots,$$

$$0 = \mu_1 < \mu_2 \leqq \mu_3 \leqq \mu_4 \leqq \cdots.$$

Polya then proves the theorem:

Let A stand for the area of D. If D is a plane-covering domain, then

$$(3) \qquad \lambda_k \geqq 4\pi k A^{-1}, \qquad k = 1, 2, 3, \cdots.$$

If D is a regularly plane-covering domain, then

$$(4) \qquad \mu_k \leqq 4\pi(k-1)A^{-1}, \qquad k = 1, 2, 3, \cdots.$$

D is said to be a plane-covering domain if an infinity of domains congruent to D, admitting congruence by symmetry, cover the whole plane without gaps and without overlapping. Concerning regularly plane-covering domains, we

need only know that the definition of a regularly plane-covering domain is more restrictive than that of a plane-covering domain.

Polya's proof of (4) is more complicated than his relatively simple proof of (3). We shall give here a proof of (4) that is analogous to and as simple as Polya's proof of (3). Furthermore, we shall prove (4) for plane-covering domains. Our method of proof of (4) is different from Polya's in that we relate the μ_k to the λ_k and use the relation (11) which concerns the λ_k. Polya works only with the μ_k, for which no relation analogous to (11) holds.

Proof of theorem. Besides the plane-covering domain D we shall consider the domains $D^{(i)}, i=1, 2, 3$. As μ_k, λ_k, A are related to D, so are $\mu_k^{(i)}, \lambda_k^{(i)}, A^{(i)}$ related to $D^{(i)}, i=1, 2, 3$.

We define $D^{(1)}$ to be a square with unit side, and so

$$A^{(1)} = 1, \tag{5}$$

$$\lim_{k\to\infty} \lambda_k^{(1)} k^{-1} = 4\pi. \tag{6}$$

The relation (6) is well known and is easily obtainable from the explicit form of $\lambda_k^{(1)}$ (see [2], pp. 429–431).

We define $D^{(2)}$ to be similar to D, so that the length of any line in $D^{(2)}$ is to the corresponding length in D as h is to 1, and so

$$A^{(2)} = h^2 A, \tag{7}$$

$$\mu_k^{(2)} = h^{-2} \mu_k. \tag{8}$$

The relation (8) is easily derived from the differential equation. $D^{(2)}$ is, like D, a plane-covering domain. Without leaving gaps, we fill the plane with non-overlapping domains congruent to $D^{(2)}$. Let n be the number of those among these congruent domains that intersect $D^{(1)}$, so that these n domains cover $D^{(1)}$. Let us define $D^{(3)}$ to be the union of these n domains; as $h\to 0$, $n\to\infty$, and so

$$\lim_{n\to\infty} nA^{(2)} = A^{(1)}, \tag{9}$$

$$D^{(3)} \supset D^{(1)}. \tag{10}$$

Relation (10) implies (see [2], p. 409) that

$$\lambda_k^{(3)} \leqq \lambda_k^{(1)}. \tag{11}$$

Let ${}^*\mu_k^{(3)}$ be the kth number, arranged in order of increasing magnitude, in the combined set of eigenvalues belonging to the n subdomains congruent to $D^{(2)}$ whose union is $D^{(3)}$. Then (see [2], p. 409)

$${}^*\mu_k^{(3)} \leqq \mu_k^{(3)}. \tag{12}$$

Furthermore, we have that (see [2], p. 410)

$$\mu_k^{(3)} \leqq \lambda_k^{(3)}. \tag{13}$$

Relations (11), (12), and (13) are all simple consequences of the maximum-minimum characterization of the eigenvalues.

By definition,

$$\overset{*}{\mu}{}^{(3)}_{(k-1)n+i} = \mu_k^{(2)}, \qquad i = 1, 2, \cdots, n. \tag{14}$$

Using (8), (14), (12), (13), and (11) we have

$$h^{-2}\mu_k \leqq \lambda^{(1)}_{(k-1)n+i}, \qquad i = 1, 2, \cdots, n. \tag{15}$$

In order to get the best upper bound for μ_k we pick $i=1$, and so

$$\mu_k \leqq h^2\lambda^{(1)}_{(k-1)n+1}. \tag{16}$$

We can rewrite this as

$$\mu_k \leqq [(k-1)n+1]h^2[\lambda^{(1)}_{(k-1)n+1}][(k-1)n+1]^{-1}. \tag{17}$$

Using (5), (7), and (9) we have

$$\lim_{n\to\infty} nh^2 = A^{-1}. \tag{18}$$

Letting $h\to 0$, $n\to\infty$, and using (6), (17), and (18) we have the resulting

THEOREM. *Let D be a plane-covering domain; then*

$$\mu_k \leqq 4\pi(k-1)A^{-1}, \qquad k = 1, 2, 3, \cdots.$$

This note was written while the author was supported by a National Science Foundation Undergraduate Research Program.

References

1. G. Polya, On the eigenvalues of vibrating membranes, Proc. London Math. Soc., 11 (1961) 419–433.

2. R. Courant and D. Hilbert, Methods of mathematical physics, I, Interscience, New York, 1953.

A GROUP-THEORETIC CHARACTERISATION OF THE PRIMES

V. O. S. OLUNLOYO AND A. D. WEISS, University of Ibadan, Nigeria

In a previous paper [1], we defined an admissible number as one whose totitives form a complete system of residues mod $\phi(m)$, and we proved that 15 is the only composite admissible number. This yielded an arithmetical near-characterisation of the primes. In this paper, we obtain first an alternative characterisation of admissibility in terms of groups of residue classes. Combining this with the well-known group-theoretic formulation of the existence of prim-

itive roots, we obtain an elementary group theoretic characterisation of the primes.

An alternative characterisation of admissibility. Let T_m be the set of totitives of a fixed natural number m.

$$T_m = \{t\colon 1 \leqq t < m,\ (t, m) = 1\}.$$

Consider R_m, the set of distinct residue classes mod $\phi(m)$ which are represented in T_m. For which m does R_m form a group under addition mod $\phi(m)$? It is obvious that every admissible number has the property. In this case R_m is of maximal order $\phi(m)$. It is perhaps conceivable that even if R_m is not of maximal order, it could yet form a group, viz. some subgroup of the cyclic group of order $\phi(m)$. That this is not so, is the content of the theorems which follow. Thus we obtain a characterisation of admissibility in terms of group notions.

THEOREM 1. *If R_m is a group under addition* mod $\phi(m)$ *then the order of R_m is maximal i.e.* $\phi(m)$.

We state two simple lemmas.

LEMMA I. *For any n, the residue classes* mod *n form a cyclic group under addition* mod *n with Class* 1 *as generator.*

LEMMA II. *If H is a subgroup of a cyclic group C and H contains a generator of C, then $H=C$.*

Proof. Let K_m denote the cyclic group of order $\phi(m)$. If R_m is a group under addition mod $\phi(m)$ then it is a subgroup of K_m. Since $(1, m)=1$ for any m, Class 1 is always in R_m and the theorem follows from Lemmas I and II.

THEOREM 2. *The necessary and sufficient condition for R_m to form a group under addition* mod $\phi(m)$ *is that m be admissible.*

Proof. If m is admissible, i.e. if R_m is complete, by Lemma I it forms a group. Conversely if R_m forms a group, Theorem 1 implies that R_m contains $\phi(m)$ distinct classes mod $\phi(m)$, viz., m is admissible.

COROLLARY 2.1. *R_m is a (cyclic) group under addition* mod $\phi(m)$ *if and only if m is either* 15 *or a prime.*

This follows from [1].

A characterisation of the primes. It is well known that the congruence classes prime to m form a group G_m under multiplication mod m and that G_m is cyclic if and only if m has a primitive root, e.g. [2]. Also m has a primitive root if and only if $m=2, 4, p^\alpha$ or $2p^\alpha$, where p is an odd prime and α a natural number. Combining these results with Theorem 2, we get

THEOREM 3. *The totitives of m form a (cyclic) group under addition* mod $\phi(m)$ *and a cyclic group under multiplication* mod *m if and only if m is prime. This pair of group structures thus serves to characterise the primes. The two groups, being*

cyclic of the same order, are isomorphic. We are thus enabled to restate the result slightly differently thus:

The totitives of m form a group under addition mod $\phi(m)$ *isomorphic with the multiplicative group* mod m *if and only if m is prime.*

References

1. V. O. S. Olunloyo and A. D. Weiss, On the completeness of the totitives of a natural number, this MONTHLY, 73 (1966) 490–494.

2. J. Hunter, Number Theory, Interscience, New York, Edinburgh and London, 1964, p. 81.

A NOTE ON GENERALIZED EQUIVALENCE RELATIONS

H. E. PICKETT, California Research Corporation, Richmond, California

J. Hartmanis [1 and 2] has introduced the notion of a partition of type n, for sets having at least n distinct elements, by means of the following definition.

A partition of X of type n is a family $\mathcal{P}_n$ of subsets of X satisfying (1) *each member of $\mathcal{P}_n$ has at least n elements, and* (2) *each n distinct elements of X belong to exactly one member of $\mathcal{P}_n$.*

Partitions of type 1 are the usual partitions of a set, and partitions of type 2 are *incidence* geometries; i.e., families of lines such that each line has at least two points and any two distinct points belong to exactly one line.

It is a well-known fact that partitions of X of type 1 correspond in a one-one way with the families of equivalence relations on X. It is the purpose of this note to demonstrate an analogous family of relations for partitions of type n. These we call generalized equivalence relations or, more specifically, equivalence relations of type n.

Let $\mathcal{P}_n$ be a partition of X of type n. Set $E_n = \{(x_1, \cdots, x_{n+1}) \mid$ for some $P \in \mathcal{P}_n$, each $x_i \in P\}$. E_n is a relation of rank $n+1$ on X, and it is easy to check that E_n satisfies the following three axioms, where we are using $\mathfrak{S}_{n+1}$ to denote the set of permutations on $\{1, 2, \cdots, n+1\}$.

E1_n. $\Lambda x_1 \cdots \Lambda x_n[(x_1, \cdots, x_n, x_1) \in E_n]$.

E2_n. $\Lambda x_1 \cdots \Lambda x_{n+1}[(x_1, \cdots, x_{n+1}) \in E_n \rightarrow \Lambda \pi \in \mathfrak{S}_{n+1}[(x_{\pi(1)}, \cdots, x_{\pi(n+1)}) \in E_n]]$.

E3_n. $\Lambda x_0 \Lambda x_1 \cdots \Lambda x_n \Lambda x_{n+1}[x_i \neq x_j$, for $i \neq j$, $i, j = 1, \cdots, n$, $\wedge (x_0, x_1, \cdots, x_n) \in E_n \wedge (x_1, \cdots, x_n, x_{n+1}) \in E_n \rightarrow (x_0, \cdots, x_{n-1}, x_{n+1}) \in E_n]$.

For partitions of type 1 these axioms reduce to

E1_1. $\Lambda x[(x, x) \in E_1]$.

E2_1. $\Lambda x \Lambda y[(x, y) \in E_1 \rightarrow (y, x) \in E_1]$.

E3_1. $\Lambda x \Lambda y \Lambda z[(x, y) \in E_1 \wedge (y, z) \in E_1 \rightarrow (x, z) \in E_1]$.

These are, of course, the conditions defining equivalence relations.

Now let X be a set with at least n elements and E_n a relation on X of rank $n+1$ satisfying E1_n, E2_n, and E3_n. For each set of n distinct elements $x_1, \cdots, x_n$

in X let $P\{x_1, \cdots, x_n\} = \{u | (x_1, \cdots, x_n, u) \in E_n\}$, and let $\mathcal{P}_n$ be the family of all $P\{x_1, \cdots, x_n\}$. From axioms $E1_n$ and $E2_n$ it follows that each member of $\mathcal{P}_n$ has at least n elements. Suppose $a_1 \in P\{x_1, \cdots, x_n\}$ and $a_1 \neq x_i$, $i=1, \cdots, n$. Also suppose u is any member of $P\{x_1, \cdots, x_n\}$. Using $E2_n$ freely, we have $(a_1, x_r, \cdots, x_1) \in E_n$ and $(x_n, \cdots, x_1, u) \in E_n$. By $E3_n$ we conclude that $(a_1, x_n, \cdots, x_2, u) \in E_n$ or that $P\{x_1, \cdots, x_n\} \subseteq P\{a_1, x_2, \cdots, x_n\}$. In a similar way, $P\{a_1, x_2, \cdots, x_r\} \subseteq P\{x_1, \cdots, x_n\}$. By a finite induction, if $a_1, \cdots, a_n$ are any n distinct elements in $P\{x_1, \cdots, x_n\}$, then $P\{x_1, \cdots, x_n\} = P\{a_1, \cdots, a_n\}$. If $a_1, \cdots, a_n$ also belong to $P\{y_1, \cdots, y_n\}$, then $P\{x_1, \cdots, x_n\} = P\{y_1, \cdots, y_n\}$. Hence each n distinct elements of X belong to exactly one member of $\mathcal{P}_n$. It is clear that if E_n arose from a partition $\mathcal{P}_n'$, the partition recovered in this way is identical to the original.

If a relation E_n satisfies $E1_n$ and $E2_n$ on a set of fewer than n elements, one sees that $E_n = X^{n+1}$. It is natural to take the partition arising from E_n as that having X as its only member. This agrees with the convention Hartmanis makes to extend the definition of partition of type n to sets with fewer than n elements.

References

1. J. Hartmanis, Generalized partitions and lattice embedding theorems, Proc. of Symposia in Pure Mathematics, Vol. II, Lattice theory, Amer. Math. Soc., (1961) 22–30.

2. ———, Lattice theory of generalized partitions, Canad. J. Math., 11 (1959) 97–106.

SETS OF INTEGERS UNDER ADDITION

R. E. Dressler, student, University of Rochester

In this note we show that sets of positive integers, closed under finite sums, may be characterized by divisibility properties. Corresponding results for sets of arbitrary integers follow trivially.

DEFINITION. *If* $S = \{s_1, s_2, \cdots\}$ *is a set of integers, we define* g.c.d. $(S) = (s_1, s_2, \cdots)$.

THEOREM. *If S is a set of positive integers, closed under finite sums, a necessary and sufficient condition that* g.c.d. $(S) > 1$ *is* $(u, v) > 1$ *for all* $u, v \in S$.

Proof. Necessity is trivial. We prove sufficiency by contradiction. Suppose g.c.d.$(S) = 1$. Order S by size, $S = \{s_1, s_2, \cdots\}$. Then, by the hypothesis, there exists a smallest integer $n > 2$ such that $(s_1, s_2, \cdots, s_n) = 1$. Let $(s_1, s_2, \cdots, s_{n-1}) = b > 1$ so that $(s_n, b) = 1$. We can then find integers c_i such that $b = \sum_{i=1}^{n-1} c_i s_i$. But,

$$b + s_n \sum_{i=1}^{n-1} |c_i| s_i = \sum_{i=1}^{n-1} (|c_i| s_n + c_i) s_i \in S \quad \text{and} \quad \left(b + s_n \sum_{i=1}^{n-1} |c_i| s_i, s_n\right) = 1$$

which contradicts the hypothesis that $(u, v) > 1$ for all $u, v \in S$. Thus if $(u, v) > 1$ for all $u, v \in S$ then g.c.d. $(S) > 1$.

As a corollary, we have the following interesting result.

DEFINITION. *A set T is a tail if and only if there exists a positive integer t, such that T contains exactly those integers $\geqq t$.*

THEOREM. *If S is a set of positive integers containing no tail and if S is closed under finite sums, then* g.c.d. $(S)>1$.

Proof. It is a well-known theorem (see [1]) that if x and y are positive integers with $(x, y)=1$ and if A is the set of all integers of the form $ax+by$, where a and b range over the set of positive integers, then A contains every integer greater than xy. Thus, since S does not contain a tail, $(a, b)>1$ for all $a, b \in S$. This implies that g.c.d.$(S)>1$.

For completeness, we include a proof of the above quoted result.

Proof. We can clearly find nonzero integers c and d such that $cx+dy=n$, for any given positive integer n. If $d<0$, let rx be the least positive multiple of x such that $d+rx>0$. Then, $(c-(r-1)y)x+(d+rx)y=n+xy$ and $(c-(r-1)y)x+xy \geqq n+xy$. Thus, $c-(r-1)y \geqq n/x>0$, and $c-(r-1)>0$. Since A is closed under finite sums, $n+xy \in A$, for every positive integer n. Thus, A contains every integer $>xy$. If $c<0$, a similar argument holds. Finally, if $c, d>0$, it follows easily that $n+xy \in A$.

The author wishes to thank the referee for his helpful comments.

Reference

1. I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, Wiley, New York and London, 1960, p. 97, problem 5.

A MEAN VALUE THEOREM—AN EXTENSION

T. V. LAKSHMINARASIMHAN, Madras Christian College, Tambaram, Madras, India

Let $f(x)$ be defined in the interval $[a, b]$ and differentiable in the interval. Further let $f'(a)=f'(b)$. Then there is a point ξ in (a, b) such that

$$\frac{f(\xi)-f(a)}{\xi-a}=f'(\xi).$$

The above result is a mean value theorem due to Flett ([1], p. 121; [2], p. 39).

In this note we obtain an extension of the above result.

THEOREM. *Let $f(x)$, defined in the interval $[a, b]$, be continuous in $[a, b]$ and differentiable at $x=a$ and $x=b$. Further let the four Dini derivates ([1]) $f^+(x)$, $f_+(x)$, $f^-(x)$, $f_-(x)$ exist and be finite in (a, b). Then if $f'(a)=f'(b)$ there exists a ξ_1, $a<\xi_1<b$, such that*

$$f^+(\xi_1) \leqq \frac{f(\xi_1)-f(a)}{\xi_1-a} \leqq f_-(\xi_1) \tag{1}$$

COMMENTS ON INVOLUTORY MATRICES

J. H. HODGES, University of Colorado and University of California, Berkeley

In a note in this MONTHLY a few years ago [vol. 63 (1956) 704–709], Nisar A. Khan proved some elementary theorems concerning involutory matrices (matrices A such that $A^2=I$, the identity matrix) with real (and in particular, rational integral) elements.

It is perhaps worth noting that

1. *All of the following of his results (and proofs given) hold for any involutory matrix A of order n over any field not of characteristic 2, with $r=$ rank $(I+A)$ and $k=$ any positive integer*:

THEOREM 3. *A is similar to* $\operatorname{diag}(I_t, -I_{n-t})$ *for some t, $0\leqq t\leqq n$, depending on A.*

THEOREM 4. *$[(I\pm A)/2]^k$ are idempotent.*

THEOREM 5. *$[(I+A)/2]^k$ and $[(I-A)/2]^k$ are orthogonal to each other.*

THEOREM 6. *rank $(I+A)^k+$rank $(I-A)^k=n$.*

THEOREM 7. *A has characteristic polynomial $|xI-A|=(x-1)^r(x+1)^{n-r}$.*

2. *The following partial analogs of his results (and proofs given) hold for any involutory matrix A over any finite field of $q=p^f$ elements (including trivially the case where $p=2$).*

Theorems 7′, 8′, 9′. $\{2(I\pm A)\}^k\mp 2A=2I$ *for $k=q$ or $2q-1$.*

This note was written under the support of NSF Science Faculty Fellowship 64203.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

A THEOREM ON POSITIVE QUARTIC FORMS

M. L. BANDY, IBM San Jose Research Laboratory

A fourth order partial differential expression

$$LU = \sum_{ijkl=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} a_{ijkl}(x_1, \cdots, x_n) \frac{\partial^2 U(x_1, \cdots, x_n)}{\partial x_k \partial x_l}$$

is said to be elliptic in a domain D in R_n if the quartic form

$$F \equiv \sum_{ijkl=1}^{n} a_{ijkl}\eta_i\eta_j\eta_k\eta_l$$

is positive definite in the interior of D. One difficulty with this definition arises when studying the properties of the expression $\int_D \bar{U} L U dx_1 \cdots dx_n$. Two applications of Green's Theorem yield

$$\int_D \bar{U} L U dx = \int_D \sum_{ijkl=1}^{n} a_{ijkl} \frac{\partial^2 \bar{U}}{\partial x_i \partial x_j} \frac{\partial^2 U}{\partial x_k \partial x_l} dx + \text{boundary terms.}$$

This focuses attention on the quadratic form

$$\hat{Q} = \sum_{ijkl=1}^{n}{}' a_{ijkl} \bar{\eta}_{ij} \eta_{kl}.$$

The definition of ellipticity gives no direct information on $\hat{Q}$. In the case $n=2$, however, V. K. Zakharov in [1] stated (but did not prove) a theorem in which the positive definiteness of F implies the positive definiteness of a quadratic form Q derived from $\hat{Q}$.

This paper is concerned with extending and proving Zakharov's theorem.

We consider the quartic form,

$$F(\xi, \eta; x_1, x_2) \equiv a\xi^4 + 4b\xi^3\eta + 6c\xi^2\eta^2 + 4d\xi\eta^3 + e\eta^4, \tag{1}$$

where the coefficients are real functions of (x_1, x_2), smooth in some domain $D(x_1, x_2)$. When we make the substitutions $\xi^2 \to w$ $\xi\eta \to y$, $\eta^2 \to z$, the corresponding quadratic form is

$$\hat{Q}(w, y, z; x_1, x_2) \equiv aw^2 + 4bwy + 4cy^2 + 4dyz + ez^2 + 2cwz. \tag{2}$$

THEOREM A. *Let $F(\xi, \eta; x_1, x_2)$ be given in $\bar{D}$, the closure of $D(x_1, x_2)$. Then there exists a function $\lambda(x_1, x_2)$, defined in $\bar{D}$, and as smooth there as the coefficients of F, such that*

(i) *inside D, $F>0$ if and only if $Q>0$;*

(ii) *on the boundary of D, $F \geqq 0$ if and only if $Q \geqq 0$; for all real ξ, η, w, y, z such that $\xi^2+\eta^2 \neq 0$, $w^2+y^2+z^2 \neq 0$, and where*

$$Q \equiv Q(\lambda, w, y, z; x_1, x_2) \equiv \hat{Q}(w, y, z; x_1, x_2) - 4\lambda(x_1, x_2)(y^2 - wz).$$

Proof. $Q(\lambda, \xi^2, \xi\eta, \eta^2; x_1, x_2) \equiv F(\xi, \eta; x_1, x_2)$. Thus, "if" is immediate in both (i) and (ii). To prove the "only if" part, without loss of generality, we can take $\eta \neq 0$. Let

$$g(t) = (1/\eta^4)F(\xi, \eta; x_1^0, x_2^0) = at^4 + 4bt^3 + 6ct^2 + 4dt + e, \tag{3}$$

where $t=\xi/\eta$, and (x_1^0, x_2^0) is any point in the interior of D. Since $g(t)>0$ for all real t, at every interior point of D, g has no real roots. This imposes conditions (see [2]) on the coefficients of g:

1) $I^3-27J^2>0$,
2) $H \geqq 0$, or $12\, H^2 < Ia^2$,
3) $a>0$, and $e>0$,

where $H=ac-b^2$, $I=ae-4bd+3c^2$, and $J=ace+2bcd-ad^2-c^3-eb^2$. Conditions 1) and 2) may be put in the more convenient form,

1′) $I>0$ and $I\sqrt{I}+3\sqrt{3}J>0$,

2′) $H+a\sqrt{I/12}>0$.

The form $Q(\lambda, x, y, z; x_1, x_2)$ can be written as

$$(4) \qquad Q = (x, y, z)\begin{bmatrix} a & 2b & c+2\lambda \\ 2b & 4(c-\lambda) & 2d \\ c+2\lambda & 2d & e \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = X^{\mathrm{T}}QX.$$

Q is positive definite if and only if the determinants of its principal minors are positive, that is,

$$(5) \qquad \begin{aligned} &\text{(i)}\ a > 0, \\ &\text{(ii)}\ ac - b^2 - a\lambda > 0, \\ &\text{(iii)}\ 4\lambda^3 - (ae - 4bd + 3c^2)\lambda + aec + 2bcd - ad^2 - c^3 - eb^2 > 0. \end{aligned}$$

Conditions (ii) and (iii) may be rewritten as follows:

(ii′) $H-a\lambda>0$,

(iii′) $4\lambda^3-I\lambda+J>0$,

where H, I, and J are defined as above. Set $\lambda=-\sqrt{I/12}$, and it is easy to verify that conditions (i), (ii′), and (iii′) follow from conditions 2′), 1′), and 3), respectively. λ is real since $I>0$ and is certainly as smooth as the original coefficients a, b, c, d, and e.

We now consider the situation on the boundary of D. Wherever F is positive definite, the previous argument holds, so we suppose $F=0$ for some (x_1', x_2') on the boundary of D and some choice of ξ and η, $\xi^2+\eta^2\neq 0$. There are two possibilities and we must examine them both.

1) Both a and e vanish for some (x_1', x_2'). Then F has the form

$$F(\xi, \eta; x_1', x_2') = 4b\xi^3\eta + 6c\xi^2\eta^2 + 4d\xi\eta^3 \geqq 0,$$

for all ξ, η, but this is impossible unless b and d vanish also. Thus F actually has the form $F=6c\xi^2\eta^2\geqq 0$, i.e., $c(x_1', x_2')\geqq 0$. The corresponding form of Q is then $Q=4(c-\lambda)y^2+2(c+2\lambda)xz$. This will be negative for some x, y, and z, unless $c+2\lambda=0$, and $c-\lambda\geqq 0$. So we take $\lambda=-c/2$ and again the condition of the theorem is satisfied. [Note that in this case $c/2=\sqrt{I/12}$.]

2) At least one of a and e is nonzero at (x_1', x_2'). Since F is symmetric in ξ and η, we can assume $a>0$ and $e\geqq 0$. Moreover, if $\eta=0$, $F(\xi, 0; x_1', x_2')=a\xi^4\neq 0$, since $\xi^2+\eta^2\neq 0$, we can take $\eta\neq 0$, and examine the polynomial $g(t)$. The roots of $g(t)$ are finite. We are assuming that $F=0$ for some real choice of ξ and η; therefore, $g(t)$ has at least one real root. A simple geometric argument shows that the four roots of $g(t)=0$ must have the form $t=t_1\pm i\beta$, t_2, t_2, where t_1, t_2, and β are real. Thus,

$$g(t) = [t - (t_1 + i\beta)][t - (t_1 - i\beta)][t - t_2]^2$$
$$= t^4 - 2(t_1 + t_2)t^3 + (t_2^2 + 4t_1t_2 + t_1^2 + \beta^2)t^2$$
$$- 2(t_2^2t_1 + t_1^2t_2 + t_2\beta^2)t + t_2^2(t_1^2 + \beta^2).$$

Comparing coefficients we have

$$b/a = -(1/2)(t_1 + t_2),$$
$$c/a = (1/6)(t_2^2 + 4t_1t_2 + t_1^2 + \beta^2),$$
$$d/a = -(1/2)(t_1t_2^2 + t_1^2t_2 + t_2\beta^2),$$
$$e/a = t_2^2(t_1^2 + \beta^2), \quad \text{and} \quad I = (1/12)a^2[(t_1 - t_2)^2 + \beta^2]^2.$$

The conditions on $Q(\lambda, x, y, a; x_1, x_2)$ for positive semidefiniteness (which are easily obtained from the matrix expression (4)) are:

$$\text{(6)} \quad \begin{aligned} &\text{(i)}\ a \geqq 0,\\ &\text{(ii)}\ ac - a\lambda - b^2 \geqq 0,\\ &\text{(iii)}\ 4\lambda^3 - (ae - 4bd + 3c^2)\lambda + ace + 2bcd - ad^2 - c^3 - eb^2 \geqq 0,\\ &\text{(iv)}\ e \geqq 0,\\ &\text{(v)}\ c - \lambda \geqq 0,\\ &\text{(vi)}\ ec - e\lambda - d^2 \geqq 0, \quad \text{and}\\ &\text{(vii)}\ ac - (c + 2\lambda)^2 \geqq 0. \end{aligned}$$

We choose $\lambda = -\sqrt{I/12} = -(a/12)[(t_1-t_2)^2+\beta^2]$. It is now a straight-forward, if tedious, calculation to verify that conditions (6) are indeed satisfied. Furthermore, with this choice of λ, condition (6-iii) reduces to an equality. Comparison with conditions (5) shows that we cannot have positive definiteness for this choice of λ, since condition (5-iii) is violated. It remains then to demonstrate that no better choice of λ exists.

We set $\hat{\lambda} = \lambda + \epsilon$ and examine again conditions (6-ii) and (6-iii):

(i) $ac - a\hat{\lambda} - b^2 = (1/4)a^2s^2 - a\epsilon = (a^2/4)(s^2 - 4\epsilon/a)$,

(ii) $4\hat{\lambda}^3 - (ae - 4bd + 3c^3)\hat{\lambda} + ace + 2bcd - ad^2 - c^3 - eb^2)$
$= -(\epsilon^2/a^2)[(t_1-t_2)^2 + (s^2 - 4\epsilon/a)]$.

It is clear that the only choice of ϵ that will make both of these expressions nonnegative is $\epsilon = 0$.

A useful extension of Theorem A may be obtained by allowing ξ, η, x, y, and z to be complex. Set

$$F(\xi, \eta; x_1, x_2) \equiv a\xi^2\bar{\xi}^2 + 2b(\bar{\xi}^2\xi\eta + \xi^2\xi\bar{\eta})$$
$$+ c(4\,|\,\xi\eta\,|^2 + \xi^2\bar{\eta}^2 + \xi^2\eta^2) + 2d(\eta^2\xi\bar{\eta} + \bar{\eta}^2\xi\eta) + e\eta^2\bar{\eta}^2,$$

$$\hat{Q}(x, y, z; x_1, x_2) = a|x|^2 + 2b(x\bar{y} + \bar{x}y) + c(4|y|^2 + x\bar{z} + \bar{x}z) + 2d(\bar{y}z + y\bar{z}) + e|z|^2.$$

THEOREM B. *Theorem A is valid if F and $\hat{Q}$ are as given in* (1) *and* (2), *when* ξ, η, x, y, *and* z *are allowed to range through all complex numbers*, (a, b, c, d, *and* e *are all real*).

Proof. 1. If $F(\xi, \eta; x_1, x_2) \geqq 0$ for all complex ξ and η, then $F \geqq 0$, for all real ξ and η.

2. Conditions (5) are necessary and sufficient for $\hat{Q}(x, y, z; x_1, x_2) > 0$, with x, y, and z complex also.

3. Conditions (6) are necessary and sufficient for $\hat{Q} \geqq 0$ with x, y, and z complex.

Thus Theorem B is a corollary of Theorem A, and the choice $\lambda = -\sqrt{I/12}$ suffices for the complex case also.

Note. It is clear from the proof of part (i) of Theorem A that if $\lambda(x_1, x_2)$ satisfies $|\lambda + \sqrt{I/12}| < \rho(x_1, x_2)$, where $\rho > 0$ is sufficiently small, then λ will still satisfy the theorem. Thus, λ can be chosen as a C^∞ function in any compact subdomain of D, bounded away from the part of the boundary of D, where $F = 0$ for some choice of ξ, η.

Acknowledgment. The author acknowledges with gratitude the valuable assistance of Dr. M. V. Menon.

References

1. V. K. Zakharov, The first boundary problem for an elliptic type equation of fourth order, degenerating on the boundary of the domain, Doklady Nauk SSSR, 114 (1957) 694–697.

2. Samuel Barnard and J. M. Child, Higher Algebra, Macmillan, London, 1936, pp. 187–193.

PATHS OF MINIMAL LENGTH WITHIN HYPERCUBES

R. A. JACOBSON, Houghton College

The reflection principle has been used quite successfully in solving many extremum problems in plane geometry [2, 3]. A recent note [1], discussing the limitations of purely analytic methods, extends the method to 3-space and employs reflections in order to find paths of minimal length within a cube. In this note we consider still another problem, suggested by M. S. Klamkin, that of finding paths of minimal length within a hypercube. Since a sequence of reflections of hypercubes in 4-space is rather hard to visualize and impossible to draw, we find the method of analysis, employed in the preceding paper [1], is not altogether useful. Thus, we reconsider the previous problem in 3-space and develop an attack involving only constructions in 2-space. Although this technique has the disadvantage of being somewhat lengthy, it is worthy of note since it is easily extended to 4-space. Not only does the following discussion disclose a solution technique, it is also beneficial in developing one's concept of 4-space and gaining insight into some of its fascinating properties.

In the following we let F_{ij} be the face of a cube lying in the plane $x_i=j$. In order to motivate the eventual attack in 4-space, we begin with a familiar problem in 3-space (prob. 1b, [1]).

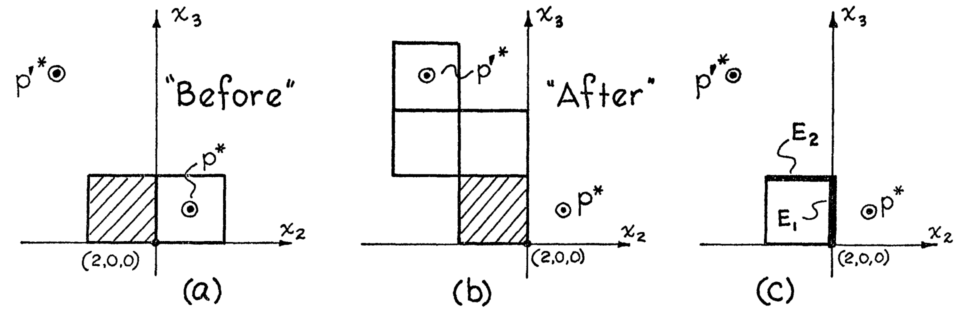


FIG. 1

Three-space problem. Let us consider the shortest closed path within the 3-cube that begins at point $p(0, 1, 1)$ on F_{10} and touches the faces F_{20}, F_{12}, F_{32}, F_{22}, F_{30} successively. We begin our analysis as before; that is, we construct a sequence of reflections through faces F_{20}, F_{12}, F_{32}, F_{22}, F_{30}. Instead of working with the resulting polyhedron, however, we consider the two polyhedra "before" and "after" reflection through face F_{12}, and project these orthogonally onto the plane $x_1=2$, see Figure 1(a) and (b). Also included are the projections, p^* and p'^*, of p and its final image p'. It is evident that all paths within the polyhedron will pass through the "common" face, shaded in both Figure 1(a) and 1(b), see Figure 1(c). In fact, the shortest such path will evidently touch either edge E_1 or E_2. We cannot find a solution directly from Figure 1(c) since it is distorted. We remedy this by "revolving" point p and p' about edges E_1, E_2 until they lie in the plane $x_1=2$, see Figure 2. "Revolving" is thought of as expressing point p in terms of its parallel and orthogonal components with respect to edge E_1 or E_2. It is clear that the minimal path intersects edge E_2 at the point

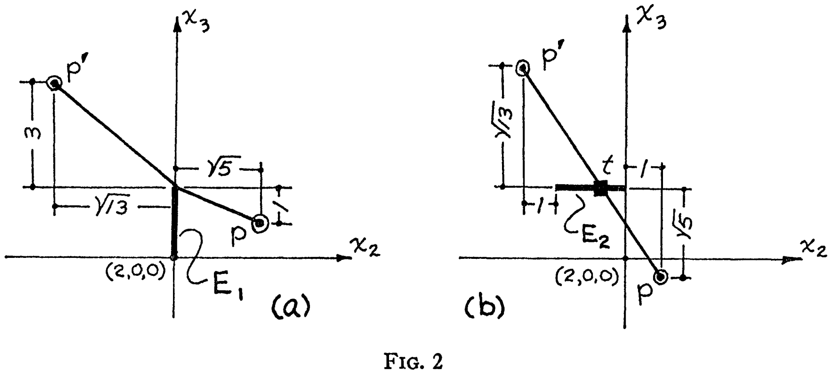


FIG. 2

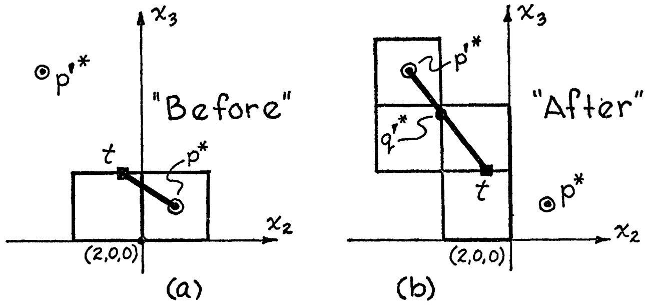


FIG. 3

[2, $(7-\sqrt{65})/2$, 2,], see Figure 2(b). The projection of the solution is easily established, see Figure 3, and the desired path can be found by inverting the projections and reflections back into the original cube.

In particular, in order to find the point q, where the path touches face F_{22}, we note that the inverse projection of the point q' is the intersection of the line connecting t and p' and the plane $x_2=-2$. Thus the reflected minimal path intersects the reflected face F_{22} at the point $[(91-\sqrt{65})/26, -2, (221-3\sqrt{65})/52]$. Reflecting this back into the original cube, we find q is the point $[(13+\sqrt{65})/26, 2, (3\sqrt{65}-13)/52]$.

A similar procedure, illustrated in the following example, will be adequate in solving the corresponding problem in 4-space.

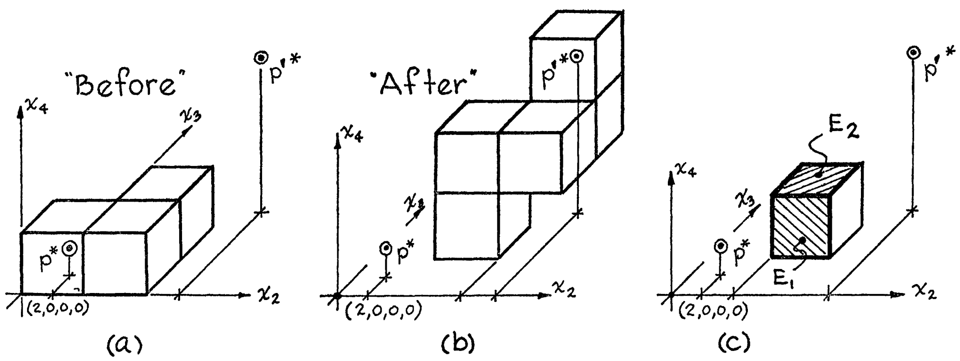


FIG. 4

Four-space problem. Let us consider the shortest closed path within the 4-cube that begins at point $p(1, 0, 0, 0)$ on F_{10} and touches faces F_{32}, F_{22}, F_{12}, F_{42}, F_{20}, F_{30}, F_{40} successively. Once again, we begin our attack with a sequence of reflections, generating some sort of polytope in 4-space. We then project the two polytopes, "before" and "after" reflection through F_{12}, onto the plane $x_1=2$, see Figure 4. Also included are the projections, p^* and p'^*, of the point p and

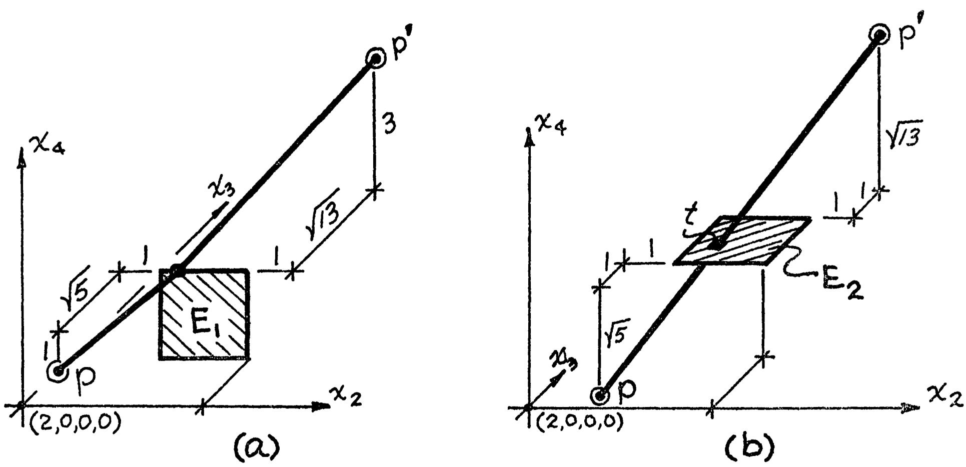


FIG. 5

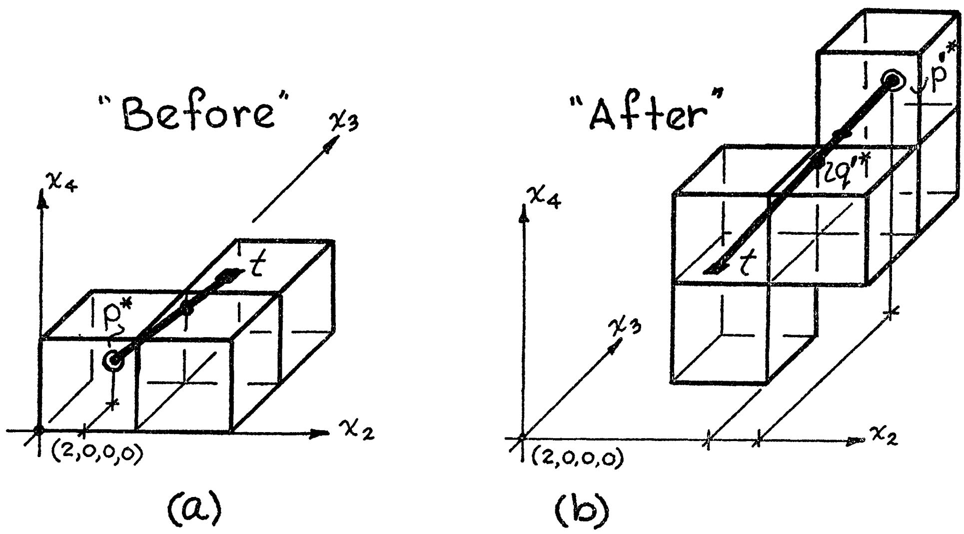


FIG. 6

its final image $p'(4, 5, 5, 5)$. It is again evident that the path of minimal length will intersect the "common" face in either edge E_1 or E_2, see Figure 4(c). Since Figure 4(c) is distorted, we first get a true picture of the situation by "revolving" points p and p' about E_1 and E_2 into the plane $x_1=2$, Figure 5. Evidently the path of minimal length intersects the "common" face at point t

$$[2, (\sqrt{65}-3)/2, (\sqrt{65}-3)/2, 2],$$

see Figure 5(b). This establishes the projection of the shortest path, see Figure 6,

and from this we can determine the desired path of minimal length.

In particular, in order to find the point q, where the path touches face F_{20}, we locate the reflected-projected point q'^*, see Figure 6(b). The intersection of the line connecting t and p' and the plane $x_2 = 4$ establishes the point q'

$$[(91 - \sqrt{65})/26, 4, 4, (221 - 3\sqrt{65})/52].$$

Reflecting q' back into the original cube, we find that q is the point

$$[(13 + \sqrt{65})/26, 0, 0, (3\sqrt{65} - 13)/52].$$

It is also evident from Figure 5(b) that the minimal path simultaneously touches both faces F_{20} and F_{30} at point q.

As in the previous paper, not all sequences will admit of a purely geometric solution (problem 1c, [1]), but a similar attack employing a function of one variable will again be sufficient to solve the corresponding problem in 4-space. Finally we point out that although the proposed solution might appear lengthy, an analysis employing only calculus would involve the almost impossible task of minimizing a function of eight radicals and twenty-one variables.

We conclude with two questions. It is evident that our attack utilized the orthogonality of the cube, as well as geometric pictures in 3-space.

(1) How would one attack a related problem for a general polytope in 4-space?

(2) How would one attack a related problem for hypercubes in n-space?

References

1. R. A. Jacobson and K. L. Yocom, Paths of minimal length within a cube, this MONTHLY, 73 (1966) 634–639.

2. N. D. Kazarinoff, Geometric Inequalities, Random House, New York, 1961, pp. 65–90.

3. F. Morley and F. V. Morley, Inversive Geometry, Chelsea, New York, 1954, pp. 35–37.

SOME DIVERGENCE PROPERTIES OF SERIES

M. E. TAYLOR, Princeton University

Given an infinite sequence of numbers whose sum diverges, we would like to know when we can pick a subsequence whose sum diverges and which is term by term less than a given sequence. It will turn out that the sequence $\{1/n\}$ plays a central role here.

Before we can proceed, we shall need a short lemma.

LEMMA. *Let $\{m_k\}$ be an increasing sequence of positive numbers, such that $m_k \to \infty$ as $k \to \infty$. Then $\sum_{k=1}^{\infty} (m_{k+1} - m_k)/m_{k+1}$ diverges.*

Proof. $\sum_{k=1}^{n-1} (m_{k+1} - m_k)/m_{k+1} \geqq (m_n - m_1)/m_n \to 1$ as $n \to \infty$. Pick n_1 so large that $(m_{n_1} - m_1)/m_{n_1} > \frac{1}{2}$. In general, given n_r, pick n_{r+1} so large that $(m_{n_{r+1}} - m_{n_r})/m_{n_{r+1}} > \frac{1}{2}$, which is possible since $(m_n - m_{n_r})/m_n \to 1$ as $n \to \infty$. Now

$$\sum_{k=1}^{n_r-1} (m_{k+1} - m_k)/m_{k+1} = \sum_{k=1}^{n_1-1} (m_{k+1} - m_k)/m_{k+1} + \sum_{k=n_1}^{n_2-1} (m_{k+1} - m_k)/m_{k+1} + \cdots$$

$$+ \sum_{k=n_{r-1}}^{n_r-1} (m_{k+1} - m_k)/m_{k+1} \geqq (m_{n_1} - m_1)/m_{n_1}$$

$$+ (m_{n_2} - m_{n_1})/m_{n_2} + \cdots + (m_{n_r} - m_{n_{r-1}})/m_{n_r} > r/2 \to \infty \quad \text{as} \quad r \to \infty.$$

Thus, since each term of $\sum_{k=1}^{\infty} (m_{k+1}-m_k)/m_{k+1}$ is positive, the series diverges.

THEOREM 1. *Let $\{t_n\}$ be a nonincreasing sequence of positive numbers, such that $t_n \to 0$ as $n \to \infty$, and $\sum_{n=1}^{\infty} t_n$ diverges. Then there is a subsequence $\{t_{n_k}\}$ such that $t_{n_k} < 1/k$ and $\sum_{k=1}^{\infty} t_{n_k}$ diverges.*

Proof. We shall extract the largest possible subsequence which is term by term less than $\{1/k\}$ and use our lemma to prove that its sum diverges. First, if there are only finitely many k such that $t_k \geqq 1/k$ then, if M is the largest of these, the sequence $\{t_{M+n}\}$ clearly satisfies the conclusions of our theorem. Thus we are justified in assuming that there are infinitely many such k.

Let n_1 be the first number such that $t_{n_1} \geqq 1/n_1$. Then $t_n \geqq 1/n_1$ for $n < n_1$. Construct a subsequence $\{t_n^1\}$ such that $t_n^1 = t_n$ if $n < n_1$.

$$\overset{1}{t}_{n_1+p} = t_{m_1+p},$$

where m_1 is the first number such that $t_{m_1} < 1/n_1$, ($p = 0, 1, 2, \cdots$). Note that $\sum_{n=1}^{\infty} t_n^1$ diverges.

As before, we can assume that $t_n^1 \geqq 1/n$ for infinitely many n. Let n_2 be the first number such that $t_{n_2}^1 \geqq 1/n_2$. Note that $n_2 > n_1$ and that $t_n^1 \geqq 1/n_2$ for $n < n_2$.

Construct a subsequence $\{t_n^2\}$ such that $t_n^2 = t_n^1$ if $n < n_2$.

$$\overset{2}{t}_{n_2+p} = \overset{1}{t}_{m_2+p},$$

where m_2 is the first number such that $t_{m_2}^1 < 1/n_2$, ($p = 0, 1, 2, \cdots$). Again $\sum_{n=1}^{\infty} t_n^2$ diverges and again we can assume that $t_n^2 \geqq 1/n$ for infinitely many n. We continue in this manner: Given a subsequence $\{t_n^s\}$, we find the first number n_{s+1} (or if we cannot, we are done) such that

$$\overset{s}{t}_{n_s+1} \geqq 1/n_{s+1},$$

and we construct a subsequence $\{t_n^{s+1}\}$ such that

$$\overset{s+1}{t_n} = \overset{s}{t_n} \text{ if } n < n_{s+1}$$

$$\overset{s+1}{t}_{n_s+1+p} = \overset{s}{t}_{m_s+1+p},$$

where m_{s+1} is the first number such that $t_{m_{s+1}}^s < 1/n_{s+1}$. Now consider the sequence $\{b_n\}$ defined by

$$b_n = t_n \text{ if } n < n_1.$$

$$b_n = \overset{1}{t_n} \text{ if } n_1 \leqq n < n_2$$

$$\cdots\cdots\cdots$$

$$b_n = \overset{k}{t_n} \text{ if } n_k \leqq n < n_{k+1}$$

$$\cdots\cdots\cdots$$

Then $\{b_n\}$ is a subsequence of $\{t_n\}$ and by construction $b_n < 1/n$. Further, we have partitioned $\{b_n\}$ into subsets, each of which, after the first, contains $n_{k+1} - n_k$ terms, such that each term is greater than or equal to $1/n_{k+1}$. Thus

$$\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{n_1-1} t_n + \sum_{n=n_1}^{\infty} b_n \geqq \sum_{n=1}^{n_1-1} t_n + \sum_{k=1}^{\infty} (n_{k+1} - n_k)/n_{k+1},$$

which diverges by our lemma. (Note: if $n_1 = 1$, use the convention that $\sum_{n=1}^{n_1-1} t_n = 0$.)

One would like to generalize the preceding result, replacing $\{1/n\}$ by an arbitrary sequence $\{a_n\}$, such that $\sum_{n=1}^{\infty} a_n$ diverges. The following theorem states that significant improvement on our first theorem is impossible.

THEOREM 2. *Given any nonincreasing sequence $\{a_n\}$ such that $na_n \to 0$ as $n \to \infty$, there is a sequence $\{t_n\}$ whose sum diverges, such that if $\{t_{n_k}\}$ is any subsequence such that $t_{n_k} < a_k$, then $\sum_{k=1}^{\infty} t_{n_k}$ converges.*

Proof. Let $a_n = e_n/n$, where $e_n \to 0$ as $n \to \infty$. Pick a subsequence

$$\{e_{m_k}\} \text{ such that } \sum_{k=1}^{\infty} e_{m_k} = R < \infty.$$

Now we want to find a sequence $\{t_n\}$ such that any subsequence, term by term less than a_n, has convergent sum. This is the case if the subsequence satisfies

$$t_{n_k} \leqq a_{m_1} \text{ if } k < m_1$$

$$t_{n_k} \leqq a_{m_2} \text{ if } m_1 \leqq k < m_2$$

$$\cdots\cdots\cdots$$

$$t_{n_k} \leqq a_{m_r} \text{ if } m_{r-1} \leqq k < m_r$$

$$\cdots\cdots\cdots$$

The m_i are defined as above, with $m_0 = 1$. Thus each a_{m_i} appears $m_i - m_{i-1}$ times. For in such a case, $\sum_{k=1}^{\infty} t_{n_k} \leqq \sum_{k=1}^{\infty} (m_k - m_{k-1}) a_{m_k}$, but

$$\sum_{k=1}^{\infty} m_k a_{m_k} = \sum_{k=1}^{\infty} e_{m_k} = R < \infty, \quad \text{and} \quad \sum_{k=1}^{\infty} m_{k-1} a_{m_k} < \sum_{k=1}^{\infty} m_k a_{m_k} < \infty.$$

Such a subsequence is forced if we let

$$t_k = a_{m_1} \text{ if } k < s_1$$
$$t_k = a_{m_2} \text{ if } s_1 \leqq k < s_2$$
$$\cdots\cdots\cdots$$
$$t_k = a_{m_r} \text{ if } s_{r-1} \leqq k < s_r$$
$$\cdots\cdots\cdots$$

where $s_1 \geqq m_1$, $s_r - s_{r-1} \geqq m_r - m_{r-1}$, and in addition s_r is sufficiently large for $\sum_{n=1}^{\infty} t_n$ to diverge. The existence of such a set s_r is obvious. Then the fact that $\{a_n\}$ is nonincreasing implies that any subsequence $\{t_{n_k}\}$ satisfies the required conditions.

An obvious corollary is that we can substitute a nonincreasing sequence $\{a_n\}$ for $\{1/n\}$ in Theorem 1 if and only if $\lim \inf_{n\to\infty} na_n > 0$.

A REMARK ON LINEAR SPACES

J. PITCAIRN, Burroughs Corp., Paoli, Pa.

Here is a quick and easy proof of Riesz's Lemma (the one which is useful in showing that local compactness implies finite dimension). It is based on the following obvious

REMARK. Let X be a linear space, A a subset, M a subspace, and λ a scalar such that $A \subset M + \lambda A$. Then for every positive integer n, $A \subset M + \lambda^n A$.

Let X be a topological linear space over a nondiscrete field with valuation. Let $\mathcal{V}$ be the neighborhood filter at 0, and Λ the set of all scalars λ such that $|\lambda| < 1$. We recall some elementary facts: a subset is *total* if it generates a dense subspace. Every element of $\mathcal{V}$ is total. If X has a total finite subset, it is finite dimensional. A subset A is *bounded* if for every $V \in \mathcal{V}$ there exists $r_V > 0$ such that $\lambda A \subset V$ whenever $|\lambda| \leqq r_V$. Every precompact subset is bounded.

LEMMA (RIESZ). *If A is bounded, and if M is a closed subspace such that for some $\lambda \in \Lambda$, $A \subset M + \lambda A$, then $A \subset M$.*

Proof. If $V \in \mathcal{V}$, take n such that $|\lambda^n| \leqq r_V$. Then $\lambda^n A \subset V$, and so by the Remark, $A \subset M + V$. Hence $A \subset M^- = M$.

COROLLARY 1. *If A is bounded and total, and if $A \subset B + \lambda A$ for some $\lambda \in \Lambda$, then B is total.*

COROLLARY 2. *If there exists a bounded total subset A such that for some $\lambda \in \Lambda$, A can be covered by a finite set of translates of λA, then X is finite dimensional.*

Proof. Take a finite F such that $A \subset F + \lambda A$. By Corollary 1, F is total.

COROLLARY 3. *If some $A \in \mathcal{V}$ is precompact, X is finite dimensional.*

Proof. A is bounded (being precompact) and total (being in $\mathcal{V}$). Take $\lambda \in \Lambda$ with $\lambda \neq 0$. Then $\lambda A \in \mathcal{V}$. Since A is precompact, it can be covered by a finite set of translates of λA. By Corollary 2, X is finite dimensional.

ON THE UNIQUENESS OF THE $*$-REPRESENTATION OF AN IDEAL

R. W. GILMER, JR. AND L. S. HUSCH, Florida State University, Tallahassee, Florida

Let R be a commutative ring with identity. An ideal A of R is *$*$-reducible* provided $A=B\cap C$ for some ideals B and C of R properly containing A such that B and C are relatively prime; that is, $B+C=R$. If A is not $*$-reducible, we say A is *$*$-irreducible*. Equivalently, A is $*$-reducible if and only if the residue class ring R/A is decomposable into a direct sum of proper subrings ([**3**], pp. 175–179, and [2], p. 42).

If R is Noetherian then each proper ideal A of R is a finite intersection of pairwise relatively prime $*$-irreducible proper ideals of R ([2], p. 41). For the sake of brevity we shall call such a representation a *$*$-representation of A*. For an arbitrary R, if A possesses a $*$-representation, such a representation is unique. For R Noetherian this result is proved in [2], pp. 38 and 41, by use of the primary decomposition theorem for ideals of R. The usual proof for the case of an arbitrary ring R uses the fact that if $A=\bigcap_{i=1}^{n} A_i$ is a $*$-representation of A, then the ring R/A is isomorphic to the direct sum of the rings R/A_i and that each of these rings is indecomposable. Then it is shown that the representation of a ring as a finite direct sum of nonzero indecomposable rings is unique. Finally, the uniqueness of this representation is used to obtain uniqueness of the $*$-representation of A ([1], p. 21).

We present in the following theorem a proof of this result using only elementary properties of the ":" operation for ideals and of relatively prime ideals, where for A and B ideals of a ring R, $A:B$ denotes the set of elements x of R such that $xb\in A$ for each $b\in B$. This proof has the additional advantage of using the same basic idea that leads to the proof of the invariance of the prime ideals belonging to the primary ideals occurring in an irredundant representation of an ideal as an intersection of primary ideals belonging to distinct prime ideals ([2], p. 35).

THEOREM. *Let A be a proper ideal of the commutative ring R with identity such that A admits a $*$-representation: $A=\bigcap_{i=1}^{r} A_i$. Then this representation is unique. That is, if $A=\bigcap_{j=1}^{s} B_j$ is another $*$-representation of A, then $r=s$ and by proper labeling, $A_i=B_i$ for all i.*

Proof. We use induction on r. If $r=1$, that is, if A is $*$-irreducible, then by definition we must have $s=1$ and $A=A_1=B_1$. We assume that for ideals possessing a $*$-representation with k components, the representation is unique. Then let $A=\bigcap_{i=1}^{k+1} A_i$, where the A_i's are pairwise relatively prime, $*$-irreducible proper ideals—that is, $A=\bigcap_{i=1}^{k+1} A_i$ is a $*$-representation of A having $k+1$ components. Let $A=\bigcap_{j=1}^{s} B_j$ be any $*$-representation of A. We let $C=A_2\cap\cdots\cap A_{k+1}$. We have $A_1+C=R$ and $C\subseteq A_i$ for $2\leqq i\leqq k+1$. Hence $A_1:C=A_1$ and $A_i:C=R$ for $2\leqq i\leqq k+1$ ([2], pp. 39–41). Therefore

$$A:C=(\cap A_i):C=\cap(A_i:C)=A_1=(\cap B_j):C=\cap(B_j:C).$$

Because $B_j+B_t=R$ for $j\neq t$ and because $B_j\subseteq B_j:C$ for all j, the ideals $B_1:C,\cdots,$ $B_s:C$ are pairwise relatively prime. Since A_1 is $*$-irreducible, $A_1=B_j:C$ for some j, say $j=1$, and $B_t:C=R$ for $2\leqq t\leqq s$. Therefore $B_1\subseteq B_1:C=A_1$ and $C\subseteq\bigcap_{t=2}^{s}B_t$. Setting $C'=\bigcap_{t=2}^{s}B_t$ we conclude, as above, that $B_1=A_v:C'$ for some v. But then $A_1\supseteq B_1\supseteq A_v$ so that $v=1$ and $B_1=A_1$. Also, $A_n:C'=R$ for $2\leqq n\leqq k+1$. It follows that $C'\subseteq\bigcap_{n=2}^{k+1}A_n=C\subseteq C'$ so $C=C'=\bigcap_{n=2}^{k+1}A_n=\bigcup_{t=2}^{s}B_t$, and these are $*$-representations of C, one of which has k components. By the inductive hypothesis, $k=s-1$ and by proper labeling $A_i=B_i$ for $2\leqq i\leqq k+1=s$. Since $A_1=B_1$, our proof is complete.

References

1. W. Krull, Idealtheorie, Chelsea, New York, 1948.
2. B. L. van der Waerden, Moderne Algebra, Springer, Berlin, 1940.
3. O. Zariski and P. Samuel, Commutative Algebra, Vol. I, Van Nostrand, Princeton, 1958.

SOLUTION OF THE QUARTIC

YANG YU-CHENG, Tainan, Taiwan, Republic of China

This paper deals with a method of solving the quartic equation

$$ax^4+bx^3+cx^2+dx+e=0. \tag{A}$$

The roots of (A) are the abscissas of the points of intersection of the curves

$$ay^2+bxy+cx^2+dx+e=0, \tag{B}$$

$$x^2-y=0. \tag{C}$$

The pencil of quadratic curves through the intersection of (B) and (C) is

$$(c-u)x^2+bxy+ay^2+dx+uy+e=0.$$

Let r be a root of

$$\Delta=\frac{1}{8}\begin{vmatrix} 2a & b & u \\ b & 2(c-u) & d \\ u & d & 2e \end{vmatrix}=0. \tag{D}$$

Then

$$(c-r)x^2+bxy+ay^2+dx+ry+e=0 \tag{E}$$

represents a pair of straight lines through the intersections of (B) and (C). Denote these lines by L_1 and L_2, their slopes by m_1 and m_2, and their point of intersection by P. Then m_1 and m_2 are the roots of

$$at^2+bt+c-r=0. \tag{F}$$

To establish this fact, transform (E) by shifting the origin to the point P. The new equation is

$$(c - r)x'^2 + bx'y' + ay'^2 + \frac{4\Delta}{4a(c - r) - b^2} = 0,$$

and the conclusion follows at once, since $\Delta = 0$.

Now the pencil of lines through the point P is

$$\text{(G)} \qquad 2(c - r)x + by + d + \lambda(bx + 2ay + r) = 0.$$

And, when $\lambda = m_1$, equation (G) represents line L_1. This comes from the fact that (G) can be written

$$y = -\frac{bm_1 + 2(c - r)}{2am_1 + b}x - \frac{rm_1 + d}{2am_1 + b}, \qquad 2am_1 + b \neq 0,$$

while (F), with $t = m_1$, can be written

$$m_1 = -\frac{bm_1 + 2(c - r)}{2am_1 + b}.$$

Likewise, when $\lambda = m_2$, equation (G) represents L_2.

Hence the roots of (A) can be found by solving (C) and (G) simultaneously. The result of this operation is

$$\text{(H)} \qquad x = \frac{\lambda}{2} \pm \frac{1}{2}\sqrt{\left(\lambda^2 - \frac{4(\lambda r + d)}{2a\lambda + b}\right)}, \qquad \lambda = m_1 \quad \text{or} \quad m_2.$$

If equation (F) has $-b/2a$ as a root, it is easy to see that (F) has a repeated root, and the roots of (A) cannot be found by the method outlined above because of the zero denominator in (H).

But if we assume that equation (E) represents two parallel straight lines, then (E) may be written

$$\text{(I)} \qquad a(y - mx - h)(y - mx - k) = 0.$$

By equating coefficients of like terms in (E) and (I), we can show that $m = -b/2a$ and that h and k are the roots of

$$at^2 + rt + e = 0.$$

Now by solving $y = x^2$ simultaneously with $y = mx + h$ we find that

$$x = \frac{\lambda}{2} \pm \frac{1}{2}\sqrt{(\lambda^2 + 4t)}, \qquad \lambda = -b/2a,\ t = h \quad \text{or} \quad k.$$

The method will be illustrated by solving the equation

$$24x^4 + 6x^3 - 13x^2 + 7x - 4 = 0.$$

We first find r by solving the equation

$$u^3 + 13u^2 + 426u + 3960 = 0.$$

By trial, $u=-10$ is a root, so we take $r=-10$. Now equation (F) becomes

$$24t^2+6t-3=0,$$

and $m_1=1/4$, $m_2=-1/2$. Upon substituting in (H), we find that the roots are

$$x=\frac{1\pm i\sqrt{15}}{8},\qquad \frac{-3\pm\sqrt{105}}{12}.$$

I am indebted to the referee, who contributed very much to the completion of the paper.

A NOTE ON THE BORDERING METHOD OF INVERTING A MATRIX

T. R. Rogge and D. D. Walling, University of Arizona

The bordering method is a computational scheme for inverting a matrix, [1, 2]. The matrix A to be inverted is the result of bordering a matrix of order $n-1$, whose inverse is already known, to obtain a matrix of order n. Formulas can then be derived which lead to the inverse of the matrix A. Thus one starts with a matrix of order 1, proceeds to a matrix of order 2, etc. In the computing of these successive inverses the value of their determinants would be of interest. The purpose of this note is to show how the determinant can be found by using quantities computed in obtaining the inverse.

The following facts are known and used by many; yet, very few textbooks set forth this relationship. It certainly could and should be used when presenting the bordering technique in the classroom.

Given the matrix A of order n, as

$$A_n=\left[\begin{array}{c:c} A_{n-1} & u_n \\ \hdashline v_n & a_{nn} \end{array}\right], \tag{1}$$

where $u_n=(a_{1n},\cdots,a_{n-1,n})^T$, and $v_n=(a_{n1},\cdots,a_{n,n-1})$. The inverse of A_n is

$$\overset{-1}{A_n}=\left[\begin{array}{c:c} \overset{-1}{A_{n-1}}+\dfrac{\overset{-1}{A_{n-1}}u_nv_n\overset{-1}{A_{n-1}}}{\alpha_n} & -\dfrac{1}{\alpha_n}\overset{-1}{A_{n-1}}u_n \\ \hdashline -\dfrac{v_n\overset{-1}{A_{n-1}}}{\alpha_n} & \dfrac{1}{\alpha_n} \end{array}\right], \tag{2}$$

where $\alpha_n=a_{nn}-v_nA_{n-1}^{-1}u_n$, $n\geqq 2$.

Let A_{n-j}^{st} denote the cofactor of the s, t element of order $n-j$. Thus,

$$\overset{-1}{A_{n-1}}=\frac{1}{|A_{n-1}|}(A_{n-2}^{ij})^T=\frac{1}{|A_{n-1}|}(A_{n-2}^{ji}), \tag{3}$$

where $|A_{n-1}| = \det(A_{n-1})$. Then

$$\alpha_n |A_{n-1}| = a_{nn}|A_{n-1}| - \sum_{p=1}^{n-1}\sum_{q=1}^{n-1} a_{nq}A_{n-2}^{pq}a_{pn} = |A_n|. \tag{4}$$

Thus,

$$\alpha_n = \frac{|A_n|}{|A_{n-1}|}, \qquad n \geqq 2. \tag{5}$$

Multiply k of these α's to obtain

$$a_{11}\prod_{p=2}^{k}\alpha_p = |A_k|, \qquad k = 2, 3, \cdots, \tag{6}$$

the desired result.

Equations (5) or (6) can be used as a check on the inversion process. It is possible in this process to obtain a singular or near singular submatrix and yet the complete matrix be non-singular. Equation (6) gives a step by step check to inform one if this has happened.

References

1. D. K. Faddeev, and V. N. Faddeeva, Computational Methods of Linear Algebra, Freeman' San Francisco, California, 1963.

2. V. N. Faddeeva, Computational Methods of Linear Algebra, Dover, New York, 1959.

A NATURAL MATRIX REPRESENTATION OF FINITE GROUPS

C. F. HARDING, Douglas Aircraft Company

It is well known that the elements of a finite group may be represented by a set of matrices obeying the same rules of combination under matrix algebra. It does not seem to have been shown, however, that such a set of matrices may be obtained directly by an operation on the group multiplication table.

THEOREM. *If, for each α_i in the finite group $G = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, we construct an $n \times n$ matrix $M(\alpha_i)$ by placing a one where α_i appears in the multiplication table of G and a zero otherwise, then the mapping $\alpha_i \rightarrow R(\alpha_i) = M(\epsilon)M(\alpha_i)$ is a faithful representation of G when ϵ is the neutral element.*

Proof. No generality is lost if $\alpha_1\alpha_2 = \alpha_3$ and thus for the theorem to be true one must have $M(\alpha_1)M(\epsilon)M(\alpha_2) = M(\alpha_3)$. So letting r_i, c_i be a particular row, column of $M(\alpha_i)$ while $\epsilon \equiv \alpha_0$ there follows

$$c_0 = r_2$$
$$c_1 = r_0.$$

So we infer that r_1 and c_2 are the correct locations in the group multiplication table.

INDEFINITE INTEGRATION BY RESIDUES II

R. P. BOAS, JR., Northwestern University, and LOWELL SCHOENFELD, Mathematics Research Center, University of Wisconsin, and Pennsylvania State University

In [2] Boas discussed the evaluation by contour integration of integrals of the form $\int_\alpha^\beta F(e^{i\theta})d\theta$, or alternatively $\int_\alpha^\beta G(\sin\theta, \cos\theta)d\theta$; his theorem should have assumed that the integrand $f(z)$ used in [2] has no singularity anywhere on the unit circumference instead of merely on the arc from $e^{i\alpha}$ to $e^{i\beta}$. In attempting to extend this work to the general case we have been led to a somewhat simpler treatment which applies to a wider class of integrals. Since the details are still rather complicated, we merely state the final result here for comparison with [2]; a full treatment appears in SIAM Review, 8 (1966) 173–183. For the closely related integral $\int_a^b F(t)dt$, we refer also to [1] (1st ed., p. 200; 2nd ed., p. 214).

THEOREM. *Let $F(z)$ be holomorphic in the extended plane except for a finite number of singularities, of which those appearing on the arc Γ: $[e^{i\alpha} \cdots e^{i\beta}]$ $(\alpha<\beta<\alpha+2\pi)$ of $|z|=1$ are simple poles not at $e^{i\alpha}$ or $e^{i\beta}$. Then the principal value of $\int_\alpha^\beta F(e^{i\theta})d\theta$ is equal to $i(R+r)$, where, with $\mu=\frac{1}{2}(\beta-\alpha)$, R is the sum of the residues of*

$$z^{-1}F(z)\log\left(e^{i\mu}\frac{z-e^{i\alpha}}{z-e^{i\beta}}\right)$$

for z in the extended plane but not on Γ, and r is the sum of the residues of

$$z^{-1}F(z)\log\left(e^{i\mu}\frac{z-e^{i\alpha}}{e^{i\beta}-z}\right)$$

for z on Γ; in both cases the logarithm has its principal value.

References

1. H. Behnke and F. Sommer, Theorie der analytischen Funktionen einer komplexen Veränderlichen, Springer, Berlin-Göttingen-Heidelberg, 1st. ed., 1955; 2nd ed., 1962.
2. R. P. Boas, Jr., Indefinite integration by residues, this MONTHLY, 71 (1964) 298–300.

THE OLD OAKEN CALCULUS PROBLEM

How dear to my heart are cylindrical wedges,
when fond recollection presents them once more,
and boxes from tin by upturning the edges,
and ships landing passengers where on the shore.
The ladder that slid in its slanting projection,
the beam in the corridor rounding the ell,
but rarest of all in that antique collection
the leaky old bucket that hung in the well—
the creaky old bucket, the squeaky old bucket,
the leaky old bucket that hung in the well.

KATHARINE O'BRIEN

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and
JOHN A. BROWN, University of Delaware

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

OBJECTIONS TO SET THEORY

P. S. MARCUS, Shimer College, and Illinois Institute of Technology

The "new mathematics" could be briefly and inaccurately defined as set theory for five-year olds. In this context, set theory refers to elementary Boolean algebra rather than the theory of infinite sets created by George Cantor. Two objections to the emphasis on set theory can be made, and the purpose of this note is to defend one of these objections as correct.

A naive objection to set theory in grammar school could be made on the grounds that it is not necessary. The layman may honestly feel that set manipulations are designed to replace useful but unglamourous arithmetic drill with useless but impressive sounding modern jargon. A proper answer to such objections is that mastery of arithmetic is not an adequate goal for learning about mathematics. The key concept which the beginning student must master is the concept of *variable*. The whole point of the "new mathematics" is that the right way to understand a variable is as a symbol that can be replaced by elements of a given set. It is the mastery of variables, both in algebra and geometry, which is the main benefit of introducing set theory.

But a second objection can be made to set theory on the grounds that it is not sufficient. The teacher who has taught his students to associate statements about x with solution sets may feel that he has fought the good fight and is entitled to rest on his laurels. This is not, however, the last word on variables. For example, after their "new math," students are going to be told that when $y=f(x)$, then $dy=f'(x)dx$, where dy is a function of two variables and dx is itself a variable. But if $x=g(t)$, then $dx=g'(t)dt$, and dx is no longer a variable but is now a function of two variables. If the student asks for a simple rule to tell him whether a given differential such as ds is to be thought of as a variable or as a function, he is not likely to receive a satisfactory answer. A little later, the student in a given situation may be told that $\partial y/\partial x=0$. Knowing that a letter used for a variable, be it r or y, does not matter at all since a variable is only an empty symbol acting as a place-holder for elements of a set, he may expect that $\partial r/\partial x=0$ also. On being told that $\partial r/\partial x=x/\sqrt{(x^2+y^2)}$ because r "depends" on x whereas y is "independent" of x, he may desire an adequate definition of "dependent variable" and "independent variable." No such definition will be found in his text.

The well-motivated student will be able to grasp the ideas behind these procedures without trouble, but the student who is either not exceptionally

well-motivated or is excessively conscientious about logical rigor may be repelled and even harmed by inconsistencies in the standard notation. What is needed for such a student is to be found in the thorough analysis of the concept "variable" which has been made by Karl Menger.

If one may be allowed to summarize Menger's position, the key to understanding variables as they are actually used in mathematics is in a methodological principle which is the polar opposite of Occam's razor. According to Occam's razor, one should not use many different concepts and explanations which all serve the same purpose. According to Menger, one should not overload the same single concept with many different mutually inconsistent meanings which all serve different purposes.

Menger points out that variables in mathematics have historically been confused with functions, indeterminates, "arbitrary constants," unknowns, coordinates, parameters, and physical quantities such as time and pressure. The beginner has always had to elucidate meaning from context. Whenever notation has several different meanings, depending on context or interpretation, some confusion and misinterpretation is inevitable. It is true that many, perhaps most mathematicians do not regard notational inconsistencies as presenting any real problem to the student. Be that as it may, the set theory of the new mathematics is not sufficient to clear up these possible confusions and any teacher who has been troubled by them may profit by an awareness of Menger's work in this area.

Reference

1. K. Menger, Calculus, a Modern Approach, Boston 1955.

A TEACH-TEST PROCEDURE FOR OBTAINING MEASURES OF MATHEMATICAL APTITUDE

R. T. HEIMER, Pennsylvania State University

Introduction. Most universities administer a wide variety of achievement tests to entering freshmen; a primary purpose of such a testing program is to identify those areas in which a student is likely to succeed and, in turn, to provide a basis for effective counseling. In mathematics, however, evidence exists [1] which indicates that instruments currently available for the purpose of predicting success in college-level mathematics courses have serious defects.

Unfortunately, at least in a measurement sense, college-bound students have received a wide variety of prior training in mathematics with varied degrees of understanding and achievement. In part, this variability is due to differences in student motivation and ability. Another source of variability, however, is due to the type, extent, regularity, and value of previous mathematical education. It is an obvious fact that achievement examinations cannot distinguish between students who did not learn material which was presented to them in a satisfactory manner and those who did not learn the material because it was presented poorly or not at all.

In view of the situation discussed above, the writer proposed the development of a new procedure of assessment, one which would provide a measure of an individual's capacity to *learn* mathematics under carefully controlled conditions. It was hypothesized that such a measure could make a unique and significant contribution to the prediction of outcomes in mathematics education. The purpose of this paper is to describe one such procedure of assessment and to examine its potential on the basis of evidence provided by a pilot study and two subsequent parallel studies.

Description of the materials. The task of obtaining a measure of a person's ability to learn mathematics led to the development of a packet of materials which, hereinafter, will be referred to as a *Teach-Test* (T-T) package; as implied by its name, it consists of an instructional unit as well as examinations. More specifically, the T-T package contains a *pretest*, an *instructional unit*, and a *posttest*. The instructional portion of the package was designed for self-instruction, the method of programming being extremely unconventional. The unit is self-contained and deals with topics of an abstract nature which are presumed to be out of the realm of the students' past experience. Accordingly, the pretest was designed to measure the extent to which the students were familiar with the content or methods of the instructional unit prior to its administration. The posttest was designed to measure the extent to which the students

(a) comprehended the content of the unit,

(b) were able to discover new (untaught) relationships, and

(c) were able to provide sound arguments in support of conjectures about untaught relationships.

In summary, the T-T package deals with the problem of obtaining a measure of an individual's mathematical aptitude (i.e., ability to learn mathematics) rather than of measuring what he already knows.

Studies of the Teach-Test package. The original study of the T-T package consisted of administering the package to a random sample of 106 entering freshmen (at The Florida State University) during the first week of the Fall Trimester, 1964. The objective was to collect data on course performance (grades) during the first two trimesters and then to ascertain the contribution, if any, that the T-T package could make in the direction of more effective prediction of success in college-level mathematics courses.

The data were collected as planned but, generally, results proved to be inconclusive. Several noteworthy items of information did come to light, however. All entering freshmen at the FSU are required to take the ETS Cooperative Mathematics Tests (CMT), Algebra I and Algebra II. Performance on these tests weighs heavily in the determination of the first (and perhaps only) mathematics course to be taken by a student. Consequently, these scores were available for all of the experimental subjects. A scattergram was constructed to show the relationship between students' T-T scores and their total scores on the

CMT. Observation of this scattergram revealed the fact that, for the given sample, high T-T scores corresponded to high CMT scores, *but not conversely.* As a matter of fact, the range of performance on T-T for the high achievers (those scoring at least 60/80, about 40 percent of the total group) was practically identical to that of the total group (T-T scores of 15–95/100 as opposed to 5–95/100).

The fact that T-T apparently was discriminating among students at the higher levels prompted two parallel and more important investigations during the summer of 1965. These studies involved the students in two NSF Secondary Science Training Programs in mathematics, one conducted at the Florida State University under the direction of the writer and the other conducted at the Florida A & M University under the direction of Dr. Israel Glover.

It is a well-known fact among directors of Secondary Science Training Programs in Mathematics that the students selected for participation typically appear to be quite similar in terms of the usual data that are collected. The primary criteria for selection include a high IQ; an outstanding high school record, especially in mathematics; high scores on standardized achievement tests; and outstanding teacher recommendations. If possible, it is customary to select only those applicants who meet all of the foregoing criteria. Yet, in spite of the apparent homogeneity of such groups, directors of these programs have often experienced a very wide range of student performance, where certain students frequently demonstrate, for seemingly no explainable reason, a propensity to learn mathematics which far exceeds that of their classmates. Since a very fine line must often be drawn in order to make the final selections for programs of this type, one wonders how many exceptional youngsters are turned away. The problem is compounded in the case of the so-called "limited background" programs where the participants are supposed to be selected from circumstances of limited educational opportunity. Under these conditions the standard barometers of success (mentioned earlier) must, perforce, be somewhat less valid.

It was reasoned, therefore, that if there is an urgency for finding and developing latent mathematical talent, then it seems justified to develop techniques which hold some promise for doing it. Because T-T had already manifested the characteristic of discriminating between high achievers, it was decided to study the relationship between performance on the T-T package and performance in the aforementioned summer programs, *especially in the case of those students, if any, who were judged by their respective faculties to have exceptional mathematical talent.*

The Florida State Study. The SSTP conducted at The Florida State University was a six-week, limited-background program designed for students who came from small rural high schools. Of the thirty participants involved in the program, five were chosen by the FSU Human Development Clinic as part of an independent research project and, consequently, did not meet the normal criteria for selection. An additional two participants who did not meet the stated

requirements were selected for special reasons which need not be elaborated here. The remaining twenty-three participants, all of whom fulfilled the given criteria, constitute the subjects of the research reported below.

The T-T package (described above) was administered on the first day of the program; the following day the Cooperative School and College Ability Test (SCAT), Form UA, and the Cooperative Reading Test, Reading Comprehension, Form 1, were administered. The scoring of the foregoing tests was withheld until after the completion of the summer period and, indeed, until after student performance in the program had been fully evaluated.

The summer program consisted of three regular courses: one in probability and hypothesis testing, one in computer programming and related mathematics, and one in earth science; each participant was also involved in a directed individual study program in mathematics. At the end of the summer, the instructors all met together to discuss the performance of each student. A five-point scale was employed for this purpose.

Rating	*Definition*
1	Outstanding performance; good potential for becoming a creative mathematician
2	Very good performance; good mathematical ability, but with only a fair chance. of becoming a creative mathematician.
3	Average performance; very little chance of becoming a creative mathematician.

The 4 and 5 rankings should need no explanation.

During the process of rating the students, the instructors expressed a desire to assign +'s and −'s to certain students within each category, and so these were adopted as admissible ratings. (In effect, therefore, a fourteen-point rating scale was employed: $1 \leftrightarrow 1+$; $2 \leftrightarrow 1$; $3 \leftrightarrow 2+$; $4 \leftrightarrow 2$; $5 \leftrightarrow 2-$; etc.). A general distribution of ratings is given in Table I. (Except for 1 ratings, this Table does not reveal the distributions within categories.)

TABLE I. Distribution of Ratings in The Florida State SSTP Mathematics Program

Rating	1+	1	2	3	4	5
No. of Students	2	2	5	9	2	3

A summary of the scores received by the twenty-three participants on the various examinations, together with their performance ratings, is given in Table II. For the convenience of the reader the listing is given (approximately) according to performance rating.

Perhaps the most significant result of the study can be gleaned from an examination of Table II. Observe that the two students who received 1+ ratings had Teach-Test scores above ninety, and that they were the *only* students with T-T scores in this category. Further observation of the Table reveals the fact that these students are *not* identifiable on the basis of any of the other measures for which scores are available.

TABLE II. Scores for all Florida State Mathematics Program Participants on all Basic Measures Together with Performance Rating

Student	Math Comp Rating	Teach-Test	SCAT (Form UA)			English Coop. (Form 1C)		
			Verbal	*Quant.*	*Total*	*Vocab.*	*Speed of Compreh.*	*Total*
# 1	1+	94	33	46	79	46	47	93
2	1+	91	48	38	86	53	56	109
3	1	81	37	40	77	50	54	104
4	1	75	33	43	76	47	48	95
5	2	81	34	36	70	47	52	99
6	2	80	39	33	72	48	53	101
7	2	75	35	44	79	44	48	92
8	2	60	44	41	85	52	46	98
9	2	40	47	38	85	50	56	106
10	3	83	43	34	77	52	53	105
11	3	78	42	47	89	48	37	85
12	3	66	41	29	70	48	55	103
13	3	65	31	30	61	43	37	80
14	3	64	48	38	86	49	53	102
15	3	63	35	42	77	42	50	92
16	3	54	23	34	57	37	41	78
17	3	52	38	39	77	53	50	103
18	3	48	48	32	80	48	51	99
19	4	32	28	35	63	36	34	70
20	4	24	22	34	56	31	21	52
21	5	64	33	38	71	52	46	98
22	5	55	52	36	88	57	53	110
23	5	23	28	28	56	46	28	74

To gain more information about Teach-Test, a matrix of correlations among the eight variables given in Table II was computed. These correlations are given in Table III, based on conversion to fourteen-point scale as previously described.

The correlations of interest are found in the last column. Observe that the highest correlate (among those considered) of performance in the program was the Teach-Test score, the nearest competition being SCAT Quantitative. A T-test revealed, however, that it would be untenable to assert that the difference in these correlations (.67 and .49) is significant. [One would need to use a level of significance of $\alpha \doteq 0.3$ in order to make such a claim. It should also be noted that there is considerable instability associated with correlations based on such a small sample ($N=23$).]

Correlations between performance in the summer program and scores on selected measures employed by the Human Development Clinic were also computed. They are presented in Table IV, where HP represents a special "high performance" measure which is a composite of several other measures.

Twelve multiple correlations, $r_{1x\cdot 8}$, also were computed (six using the scores

TABLE III. Intercorrelations Among all Measures Administered as a Regular Part of the Study ($N=23$)

	1 T-T	2 SCAT V	3 SCAT Q	4 SCAT T	5 CE V	6 CE SC	7 CE T	8 Performance Rating
1. T-T	—	.30	.47	.48	.43	.60	.57	.67
2. SCAT-V		—	.13	.87	.80	.72	.81	.15
3. SCAT-Q			—	.61	.16	.19	.19	.49
4. SCAT-T				—	.72	.67	.74	.37
5. CE-V					—	.71	.89	.10
6. CE-SC						—	.95	.44
7. CE-T							—	.33
8. Performance Rating								—

TABLE IV. CORRELATIONS BETWEEN SCORES ON SELECTED MEASURES ADMINISTERED BY THE HUMAN DEVELOPMENT CLINIC AND PERFORMANCE IN THE MATHEMATICS PROGRAM

WAIS-FS	.32	WAIS-P	.23	STEP-Total	.27
WAIS-V	.36	BINET IQ	.04	HP	.23

on the regularly administered measures and six using scores on the measures listed in Table IV); the greatest of these was .72, a result not significantly greater than the simple correlation $r_{18} \doteq .67$. No further analyses were attempted.

The Florida A & M Study. The study conducted at the Florida A & M University was more limited in scope than the Florida State study, but the events which transpired seem worthy of reporting.

As at FSU, the T-T package was administered to the (twenty-five) participants of their SSTP in mathematics on the first day of that program. No other tests were administered to this group as part of the present study.

At the completion of the program, the A & M faculty ranked their students in order from 1 to 25. The Teach-Test scores, according to rank, are given in Table V.

The rank-order correlation coefficient was computed and found to be .64. Once again, moreover, the package seemed to be especially effective in identifying the really exceptional students. Observe that the highest T-T score, namely 69, belonged to the top-ranked student (who, according to his instructors, "stood

TABLE V. TEACH-TEST SCORES OF PARTICIPANTS IN FAMU STUDY ACCORDING TO (RANK) PERFORMANCE

Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
T-T *Score*	69	56	50	34	58	44	28	52	30	28	35	51	29	40	43	47	29	17	52	24	22	41	16	32	19

out" from among the rest of the group). It might also be noted that, in a very crude sense, these data provide for a cross-validation of the Florida State results. That is, the results are consistent; and this fact provides some degree of confidence in the validity of the T-T instrument even though the samples are small.

Conclusions. The fact that the instructional units in T-T packages can be cast in a programmed (self-teaching) format is a guarantee of the feasibility of administering such instruments—much the same as ordinary measurement instruments are administered. Furthermore, T-T packages certainly could be developed for use at a variety of educational levels, a fact which suggests a multitude of intriguing possibilities.

Although the investigations reported above should be interpreted as nothing more than pilot studies, they seem to have provided relevant evidence in support of the merit of the Teach-Test procedure and its potential for obtaining valid measures of mathematical aptitude. There seems to be substantial promise in utilizing such a procedure for assisting in the search for latent mathematical talent—wherever it exists.

Reference

1. E. A. Coddington, Scholastic aptitude tests in mathematics, this MONTHLY, 70 (1963) 750–755.

BRIEF COMMENT

Paperbacks: The New Math, MORRIS KLINE, The New York Times Book Review, March 8, 1966.

"Civil rights, the war in Vietnam, inflation and taxes—these are minor problems compared to one which millions of American parents are now facing, that of helping their children learn the 'new mathematics.' The problem isn't made any easier by the fact that the stores are currently displaying a dozen paperback books about mathematics whose covers suggest, somewhat misleadingly, that they will provide first aid to the perplexed. As a matter of fact, only six of the books are of value to parents, and these only with qualifications."

The books reviewed include Evelyn B. Rosenthal's *Understanding the New Mathematics*, Charles M. Barker Jr., Helen Curran and Mary Metcalf's *The New Math*, Carl B. Allendoerfer's *Mathematics for Parents*, Irving Adler's *The New Mathematics*, and Jerome S. Meyer and Stuart Hanlon's *Fun with the New Math.* Besides the reviews, a brief history of the recent changes in the school curriculum, i.e. the introduction of the "new mathematics" is given.

Turning a Surface Inside Out, ANTHONY PHILLIPS, *Scientific American*, May, 214 (1966) 112–120.

An expository article, on a highly simplified level, of some work done by Stephen Smale and others in the area of differential topology—"The great mathematician David Hilbert once said that a mathematical theory should not be considered perfect until it could be explained to the first man one met in the street. Hilbert's successors have generally despaired of living up to this standard. From time to time, however, research on an advanced and inaccessible mathematical topic leads to a discovery that is intuitively attractive and can be explained without oversimplification. A striking example is Stephen Smale's theorem concerning regular maps of the sphere, published in 1959. A visualization devised by the late Arnold Shapiro of Brandeis University enables us to depict a startling consequence of Smale's theorem: it is possible, from the topologist's point of view, to turn a surface such as a sphere inside out." The article also deals with such concepts as regular curves, regular homotopy, and regular maps. The illustrations, as usual, are superb.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to M. S. Klamkin, Ford Sci. Lab., P.O. Box 2053, Dearborn, MI 48121. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before February 28, 1967.

Editorial Announcement. **We regret that we cannot be responsible for any solution to any Elementary Problem, submitted on or after October 1, 1966, unless it was directed to M. S. Klamkin at the foregoing address.**

E 1915. *Proposed by Kaidy Tan, Fukien Normal College, Foochow, China*

Construct a triangle ABC given side BC, the median AM, and the angle bisector AT (that is, given a, m_a, t_a).

E 1916. *Proposed by J. E. Lewis, Acadia University, Wolfville, N.S., Canada*

Given any positive integer n, then if both $(a, n)=1$ and $(a+2, n)=1$, we will call a and $a+2$ a *special pair*; and we define the function $\psi(n)$ as the number of special pairs less than n. Prove the

THEOREM. $\psi(n)+1=n\prod_{i=1}^{k}(1-2/p_i)$, *where* $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$ *is the canonical decomposition of* n, $(2, n)=1$;

$$\psi(m)+1=2^{\alpha-1}n\prod_{i=1}^{k}(1-2/p_i), \quad \textit{where} \quad m=2^{\alpha}n,\ (2,n)=1.$$

E 1917. *Proposed by K. O. May, University of Toronto*

The ordered pair is customarily defined in terms of the unordered pair by $(a, b)=\{\{a, b\}, \{a\}\}$. Prove that it is impossible to reverse this by defining unordered pair in terms of ordered pair.

E 1918. *Proposed by L. J. Warren and Jerry Tice, San Diego State College*

Let p be a prime larger than 3. Show that there is no positive integer k such that $3p\,|\,\sigma_k(3p)$, where $\sigma_k(n)$ is the sum of the kth powers of the divisors of n.

E 1919. *Proposed by D. R. Rao, Secunderabad, India*

Find the roots of the equation $x^6-px^4+qx^3-rx^2+s=0$, having given $p(q^2-4s)=r^2$.

E 1920. *Proposed by H. S. M. Coxeter, University of Toronto*

Find the eccentricity of an ellipse which passes through two of the cusps of its evolute.

E 1921. *Proposed by H. S. M. Coxeter, University of Toronto*

Prove that there are at most three points on or inside the unit circle such that the distance between any two of them is greater than $\sqrt{2}$.

E 1922. *Proposed by Gregory Wulczyn, Bucknell University*

Find a (least) prime which is simultaneously of each of the forms: $x_1^2+y_1^2$, $x_2^2+2y_2^2, \cdots, x_{10}^2+10y_{10}^2$.

E 1923. *Proposed by I. J. Good, Trinity College, Oxford, England*

Prove that the square of the area of a face of any polyhedron is equal to the sums of the squares of the areas of the other faces, minus twice the sum, over every pair of the other faces, of the products of their areas times the cosine of the angle between them.

E 1924. *Proposed by Mason Henderson, Montana State University*

Let A and B be unit disks, A cut by a chord into two parts A_1 and A_2. (1) Is there a square of least area which covers A_1 and A_2, where A_1 and A_2 are placed

so as not to overlap? (2) Is there a square of least area which covers B and both A_1 and A_2? (3) Replace square by rectangle in (1) and (2).

SOLUTIONS OF ELEMENTARY PROBLEMS

E 1646 [1964, 917]. *Correction by Edward Milner, NASA—Lewis Research Center, Cleveland, Ohio.* The solution as shown is incorrect.

$0 \in A'$. Hence, $A' = \{0, 2, \cdots, k+1\}$ is closed, contradicting one of the conditions of the problem.

The following is a correct solution to the problem: Let $S = \{1, 2, 3, \cdots, k+1, a, b\}$ with basis $\{\{1\}, \{1, 2\}, \{1, 2, 3\}, \cdots, \{1, 2, 3, \cdots, k+1\}, \{a, b\}\}$. Let $A = \{1, 2, 3, \cdots, k+1, a\}$ which is not closed. For $1 \leqq j \leqq k$, $A^{(j)} = \{j+1, \cdots, k+1, b\}$ if j is odd, and $A^{(j)} = \{j+1, \cdots, k+1\ a\}$ if j is even. Noting that every open set containing either a or b contains both a and b, we see that each of the sets $A, A', A'', \cdots, A^{(k)}$ is not closed.

A Solid with Certain Cross-sections Specified

E 1793 [1965, 544–545]. *Proposed by Walter Bluger, Dominion Bureau of Statistics, Ottawa, Canada*

The projection of a solid body onto three mutually perpendicular planes yields a unit disk and two areas with square boundaries. Describe the possible shapes of the bodies.

Solution by Michael Goldberg, Washington, D. C. The most obvious shape is a cylinder whose height is equal to its diameter.

A second satisfactory shape is a "double cone," fitted together base to base, with the height of each cone equal to the radius of its base. Variations of this solution may be obtained by deforming the circular cross-sections except at the midsection.

A third family of solutions arises as follows: A ruled surface may be made by straight lines from points along a circle to points along two parallel straight lines at a distance from the plane of the circle equal to a radius. The association of corresponding points may be made in an infinite number of ways. Furthermore, the surface may be varied by minor deformations without changing these views provided, only, that the straight lines at the four cardinal points of the circle are retained.

Also solved by D. M. Hancasky, M. S. Klamkin, Charles McCracken, D. C. B. Marsh, and the proposer.

A Group Consisting of nth Powers

E 1794 [1965, 545]. *Proposed by Alan Schwartz, University of Wisconsin*

Let G be a finite abelian group, multiplicatively written; and let n be the order of G. Then n is odd if and only if each element of G is a square.

I. *Solution by E. A. Fay, Naval Ordnance Test Station, China Lake, California.* The result holds without the restriction that G be abelian. Suppose first that n is odd, and let x be any element of G. Since the order of each element divides n, x is of odd order, and $x^{2k-1}=e$ for some positive integer k, where e is the identity. Then $x^{2k}=x$, so x is a square.

Suppose conversely that n is even. Since the number of elements other than e is odd, at least one of them must be its own inverse. Then, on the diagonal of the multiplication table of G, e appears at least twice, so some element, say y, must not appear at all. Thus y is not a square.

II. *Solution by Robert Patenaude, Humboldt State College.* Delete the condition that G is abelian, and generalize the problem to end: Then $(k, n)=1$ if and only if each element in G is a kth power.

Suppose $(k, n)=1$. If a typical element a in G is of order n', then $n'|n$ by Lagrange's theorem. The congruence $kx\equiv 1 \pmod{n'}$ has a solution since $(k, n')=(k, n)=1|1$, so that $(a^x)^k=a$.

Suppose instead that $(k, n)>1$; choose a prime p so that $p|(k, n)$. By Cauchy's Theorem (I. Herstein, *Topics in Algebra*, pg. 74), there exists in G an element (say b) of order p. Since $p|k$, $e=b^p=b^k$. For every element in G to be a kth power, it is necessary that the kth powers of the n elements be distinct. Since $e^k=b^k=e$, this is impossible.

Also solved by O. P. Anand, R. A. Avelsgaard, Ralph Bennett, S. K. Bhandari (India), J. L. Brown, Jr., E. O. Buchman, T. J. Burke, B. J. Cerimele, Orin Chein, J. K. Cole, József Dénes (Hungary), J. F. Dillon, W. G. Dotson, Jr., Vlaicu Drilea (Nigeria), E. W. Ewing, W. F. Feeny, N. J. Fine, David Finkel, W. B. Galvin, R. J. Gazik and R. F. Rossa (jointly), D. M. Goldschmidt, Israel Grossman, Harry Guess, J. K. Halbig, D. M. Hancasky, R. B. Hardin, Jr., R. E. Harper, L. D. Haugh, Edward Hook, Bernard Jacobson, Erwin Just, U. S. Kahlon (India), Geoffrey Kandall, Irving Katz, Omar Khayyam, Jr., M. S. Klamkin, M. L. Klasi, B. G. Klein, W. B. Laffer II, Horst-E. Lahmann (Germany), E. S. Langford, Steve Ligh, Loe Lipman, Richard Loebl, Jiang Luh, Charles McCracken, Robert McGuignan, D. C. B. Marsh, J. J. Martinez, C. J. Maxson, Y. Mayer ben-David, Jerry Metzger, William Moser, N. S. Natarajan (India), F. D. Parker, B. J. Parshall, C. B. A. Peck, J. W. Petro and E. A. Schreiner (jointly), Harsh Pittie, J. R. Porter, Donald Quiring and Elias Toubassi (jointly), Stephen Rhodes, Azriel Rosenfeld, J. P. Ruebsamen, P. J. Ryan, S. K. Sehgal, Robert Shanny, R. Sivaramakrishnan (India), D. A. Smith, R. T. Smythe, Al Somayajulu, Mike Sterling, John Stout, J. P. Tarwater, Guy Torchinelli, David Vasholz, L. J. Warren, W. C. Waterhouse, Albert Wehrly, P. Willeg, K. L. Yocom, D. A. Zave, and the proposer.

Many solvers observed that the assumption of commutativity is not needed. Kahlon, Khayyam, Lipman, Natarajan, and Petro and Schreiner proved the generalization offered in Solution II above, while Kandall, Pittie, and Vasholz obtained the intermediate result: If p is a prime, then $p \nmid o(G)$ if and only if $G=\{x^p | x\in G\}$.

Rosenfeld points out that this problem is Theorem 1 of W. R. Utz, *Square roots in groups*, this MONTHLY, 60 (1953) 185–186. Dénes calls attention to C. Pasztor and J. Dénes, *Magyar Tudományos Akademia Matematikai Fizikai tudományok Osztályanak Kozleményei*, XIII (1963) 109–118 [see Math. Rev., 29 (1965) 180] and to K. Fenchel, Math. Scand., 10 (1962) 182–188 [see Math. Rev., 27 (1964) 197].

A Consequence of Problem 4964

E 1795 [1965, 665]. *Proposed by N. D. Kazarinoff, University of Michigan*

Let $ABCDEF$ be a convex hexagon such that the perimeters of the triangles ABF, BCD, DEF, and BDF are the same. Show that the hexagon must be a triangle, that is, it must have three $180°$ angles. Compare 4964 [1962, 672].

Solution by M. S. Klamkin, Ford Scientific Laboratory, Dearborn, Michigan. Draw $\triangle GHI$, where $GB\|AC$, $HD\|CE$, and $IF\|AE$. Since the hexagon is convex it lies inside or on $\triangle GHI$. By the result of problem 4964 [1962, 672], the perimeter of $\triangle BFD$ cannot be less than each of the triangles GFB, FDI, and BDH, and if $\triangle BFD$ has the same perimeter as one of these other triangles, then all the four triangles have the same perimeter. Since $\triangle ABF \leqq \triangle GBF$ in perimeter and similarly for the other two pairs of triangles $\measuredangle ABC = \measuredangle CDE = \measuredangle EFA = \pi$, and the hexagon is a triangle.

The above result would still be valid if we replaced "perimeter" of the triangles by "area." This follows from a corresponding result to problem 4964 and is also a solved problem in this MONTHLY, 67 (1960) 479.

Also solved by Neal Felsinger, J. M. Quoniam (France), Robin Sibson (England), and the proposer.

An Asymptotic Formula

E 1796 [1965, 665]. *Proposed by Louis Comtet, Boulogne, France*

Show that, as $n \to \infty$, $s(n) = 1^n + 2^n + 3^n + \cdots + n^n$ is asymptotic to $en^n/(e-1)$.

Solution by M. S. Klamkin, Ford Scientific Laboratory, Dearborn, Michigan. Let $\nu(n) \geqq 0$ for $n = 1, 2, 3, \cdots$, and set $S(n) = \sum_{r=1}^{n} r^{\nu(n)}$. By Maclaurin's Integral Test,

$$S(n) \geqq \int_0^n x^{\nu(n)} dx \geqq S(n) - n^{\nu(n)}.$$

Consequently,

1. If $\nu(n)/n \to 0$, then $S(n) \sim n^{\nu(n)}/[1+\nu(n)]$;
2. If $n/\nu(n) \to 0$, then $S(n) \sim n^{\nu(n)}$;
3. If $\nu(n)/n \to a+>0$, then $1/a \leqq S(n) \leqq 1+1/a$.

In Case 3, we have

$$S(n)/n^{\nu(n)} = \sum_{r=0}^{n-1} \exp[\nu(n) \log (1 - 1/n)] = \sum_{r=0}^{n-1} \exp[-\nu(n)[r/n + O(r^2/n^2)]]$$

$$= \sum_{r=0}^{n-1} e^{-ar}e^{O(1)} \to \sum_{r=0}^{\infty} e^{-ar} = \frac{e^a}{e^a - 1}$$

by Tannery's Theorem [see, for example, T. J. Bromwich, *Introduction to the Theory of Infinite Series*, 2nd ed. (1925), page 136].

Also solved by R. A. Avelsgaard, W. O. Egerland, R. B. Eggleton, P. G. Engstrom, J. A. Faucher, Robert Forbes, R. F. Jackson, V. K. Krishnan (India), Norman Miller, Simeon Reich (Israel), Henry Ricardo, M. N. Sastry, Richard Sinkhorn, Robert Spira, P. D. Thomas, L. E. Ward, Sr., David Zeitlin, and the proposer. Partial solutions by Daniel Ansimov, Paul Cull, R. J. Gazik and R. F. Rossa (jointly), H. M. Gehman, Michael Goldberg, P. G. Kirmser, E. S. Langford, L. W. Mendale, C. S. Ogilvy, Walter Penney, D. L. Silverman, Al Somayajulu, and K. L. Yocom.

None of the partial solutions contained any justification that

$$\lim_{n\to\infty} \sum_{k=1}^{n} a_n(k) = \sum_{k=1}^{\infty} \lim_{n\to\infty} a_n(k).$$

Ricardo found (essentially) E 1796 as Problem 24 on page 467 of C. A. Stewart, *Advanced Calculus*, 3rd ed., Methuen, London (1951). He also calls attention to R. P. Boas, Jr., *Tannery's Theorem*, Math. Mag. 38 (1965), 66. Ward notes that E 1796 is equivalent to Example 19 on page 528 of T. J. Bromwich, *An Introduction to the Theory of Infinite Series*, 2nd ed. (1926), and Kirmser points out Exercise 15e on page 107 of K. Knopp, *Theory and Application of Infinite Series* (1928).

Another Diophantine Equation

E 1797. *Proposed by W. J. Blundon, Memorial University of Newfoundland*

Find all solutions in integers of the equation $y^2+y=x^4+x^3+x^2+x$.

Solution by D. C. B. Marsh, Colorado School of Mines. Multiply the given equation by 4 and add 1 to get

$$4x^4 + 4x^3 + 4x^2 + 4x + 1 = (2y + 1)^2.$$

For $x=-1$, $y=-1$ or 0; for $x=0$, $y=-1$ or 0; for $x=2$, $y=-6$ or 5. (For $x=1$, y is nonintegral.)

For all $x<-1$ or $x>2$, the left-hand side of the equation above is greater than $(2x^2+x)^2$ but less than $(2x^2+x+1)^2$ and cannot be an integral square for integral x, while the right-hand side is an integral square for integral y. Thus, the six solutions listed above are the only integral solutions of the given equation.

Also solved by Paul Devine, Michael Fredman, D. R. Morrison, C. B. A. Peck, Henry Ricardo, D. L. Silverman, Simon Vatriquant (Belgium), and the proposer. Partial solutions by J. N. Davies, R. B. Eggleton, G. G. Ford, D. M. Hancasky, G. F. Henry, R. F. Jackson, Sam Kravitz, R. E. Maas, Norman Miller, Sam Newman, Robin Sibson (England).

Ricardo calls attention to the closely related Problem 2784 on pg. 10 of *The Otto Dunkel Memorial Problem Book.*

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before April 30, 1967.

5420. *Proposed by T. P. Kezlan, University of Texas*

Suppose that R is a ring with no nonzero divisors of zero and with the property that every proper subring of R is finite. Prove that R is a field of prime characteristic.

5421. *Proposed by D. J. Newman, Yeshiva University*

Evaluate

$$\int_0^\infty \left| \int_0^{\pi/2} \left(1 - \frac{\sin xy}{\tan x}\right) dx \right| dy.$$

5422. *Proposed by A. A. Mullin, University of California*

Let ζ be the Riemann zeta-function restricted to the real numbers. Prove the following: $\zeta(s-1)\cdot\zeta(2s) \geqq \zeta^2(s)\cdot\zeta(2s) \geqq \zeta^2(s) \geqq \zeta(2s)$, for $s>2$. Does one have in fact strict inequalities? Which, if any, of the following inequalities hold:

$$|\zeta(s-1)\cdot\zeta(2s)| \geqq |\zeta(s)|^2 \cdot |\zeta(2s)| \geqq |\zeta(s)|^2 \geqq |\zeta(2s)|$$

for the extension of ζ to Re $s>2$?

5423. *Proposed by James Singer, Brooklyn College*

Suppose that

$$u_n = \frac{1}{2^n} \sum_{(n)} \sum_{i=1}^{n} \frac{e_i}{2^{2i-s_i}},$$

where $e_i = 0$ or 1 for all i, $s_i = e_1 + e_2 + \cdots + e_i$, and the sum $\sum_{(n)}$ is extended over all possible allowable choices of $e_1, \cdots, e_n$; find a simple expression for u_n and $\lim_{n\to\infty} u_n$.

5424. *Proposed by Van E. Wood, Battelle Memorial Institute*

The relation of Ramanujan

$$\text{(A)} \qquad \int_0^\infty \frac{e^{-yx}}{x(\pi^2 + \ln^2 x)}\, dx = e^y - \int_0^\infty \frac{y^x}{\Gamma(1+x)}\, dx, \qquad y \geqq 0$$

is readily verified by the Laplace transform method (G. H. Hardy, *Ramanujan*, University Press, Cambridge, 1940, p. 196). Ramanujan, who, according to Hardy, disliked contour integration, proved (A) indirectly by discovering a generalization, the correctness of which could be shown without using integration in the complex plane. Show that (A) can be verified directly (that is, without finding a more general integral of which it is a special case) by a method which does not involve contour integration.

5425. *Proposed by V. R. R. Uppuluri, Oak Ridge National Laboratory*

Let A and B be two real, symmetric, positive definite matrices. Show that $C = 2A(A+B)^{-1}A + 2B(B+A)^{-1}B - (A+B)$ is a nonnegative definite matrix.

5426. *Proposed by Omar Khayyam, Jr., University of California, Berkeley*

Prove or disprove: Every topological space X can be embedded in a simply-connected space X^* where X^* is obtained from X by the adjunction of one point.

5427. *Proposed by J. W. Wyman, Pasadena College*

Let X_1, X_2, X_3 be points in E^n, Euclidean n-space with $d(X_1, X_3)=d(X_1, X_2)+d(X_2, X_3)$. Let $S_{\epsilon_i}(X_i)$ denote the open sphere of radius ϵ_i and center X_i in E^n. If $\epsilon_3 < d(X_2, X_3)$, determine how small ϵ_1 and ϵ_2 must be so that

$$(*) \qquad x_1 \in S_{\epsilon_1}(X_1), \qquad x_2 \in S_{\epsilon_2}(X_2), \qquad x_1 \neq x_2 \Rightarrow x_1x_2 \cap S_{\epsilon_3}(X_3) \neq \phi.$$

Find the least upper bound of all ϵ' such that $\epsilon' = \epsilon_1 + \epsilon_2$ and (*) is met.

5428. *Proposed by P. G. Jessup, Lehigh University*

For two vector topologies T, T' on a linear space, let $T \wedge T'$ be the largest vector topology included in $T \cap T'$. Let $X^{\#}(X')$ be the set of all (continuous) linear functionals on X. For $S \subset X^{\#}$, let $\sigma(S)$ be the weak topology by S. For (X, T) a linear topological space, when is it true that $\sigma(X^{\#}) \wedge T = \sigma(X')$?

5429. *Proposed by J. M. Moser, San Diego State College*

Consider k stochastically independent sequential tests $\lambda_{1n}, \cdots, \lambda_{kn}$. An hypothesis is accepted when $\lambda_{in} < \beta$ for all i or rejected when $\lambda_{in} > \alpha$ for all i. Find $E[n]$ the expected number of trials for accepting or rejecting the null hypothesis.

SOLUTIONS OF ADVANCED PROBLEMS

Connected Topological Spaces not E^n

5214 [1964, 689]. *Proposed by E. D. Nix, Norwich, Vermont*

Exhibit, or prove the nonexistence of, an arcwise connected topological space S having more than one point, and which is not a homeomorph of any open set of any finite-dimensional Euclidean space, such that:

a) homeomorphisms between subsets of S preserve the property of having nonempty interiors;

b) if p and q are any (not necessarily distinct) points of S and if U_p and U_q are any neighborhoods respectively of p and q, then there exists a neighborhood V_q of q such that $V_q \subseteq U_q$ and V_q is homeomorphic to U_p with a homeomorphism $h: U_p \rightarrow V_q$ such that $h(p)=q$.

Solution by J. R. Isbell, Case Institute of Technology. The following example S is contractible and paracompact; the question whether a compact metric example exists is unresolved. S is the set of all paths $f: [0, r] \rightarrow E^2$ ($r \in [0, \infty)$) made up of a finite number of isometric embeddings of subintervals $[P_i, P_{i+1}]$ horizontally or vertically in E^2, without immediate retracing, with $f(0)=(0, 0)$. From each f in S we can describe four "principal rays" corresponding to the three natural continuations and ($r>0$) one way back. Define a subset of S to be open if it contains an initial segment of each principal ray emanating from each of its points. (Two paths f_1, f_2 are "close" if one is a "slight" extension of the other.)

Evidently each point f cuts $S-\{f\}$ into four components. Every two points are joined by a unique arc. S can be contracted by α_t taking each f on $[0, r]$

to $f|[0, tr]$. Every connected open set has a discrete boundary, whence (easily) S is paracompact. Every homeomorph of a letter "X" in S contains segments of the principal rays emanating from its center, whence homeomorphs of open sets are open. For homogeneity, S is a group with separately continuous multiplication, the path fg being traced as follows: trace f, then go along the translate $g+f(r)$, erasing as long as this retraces f. To map S into any neighborhood of 0, note that any function β: $S\to S$ continuous on principal rays is continuous, and proceed boldly: compress every f in S by composing its initial straight part f_0: $[0, p_1]\to E^2$ with $x\to\epsilon$ arctan x for suitable ϵ, compressing its second straight part by ϵ^* arctan, where ϵ^* depends on how far the given neighborhood extends from ϵ arctan $f(p_1)$, and so forth.

Consecutive Composite Integers

5317 [1965, 796]. *Proposed by Daniel I. A. Cohen, Princeton University*

Show that there exists a number k such that there are $n-1$ consecutive composite integers less than e^{kn} for all n.

I. *Solution by M. V. Subbarao, University of Alberta.* Let A denote the product of all primes not exceeding n. As is well known (see, e.g., Hardy & Wright, *Introduction to the Theory of Numbers*, 4th ed., Theorem 415) $A\leqq 4^n$. Clearly, $A+2, A+3, \cdots, A+n$ are all composite integers and $A+n<e^{kn}$ where k is a fixed number (independent of n) less than $1+\log 4$.

II. *Solution by K. K. Norton, University of Illinois.* Define $p(g)$ as the smallest prime following g or more consecutive composite integers. We shall show that $p(n-1)\leqq(14/13)\exp(n+3/2)$ for all $n\geqq 1$.

Let p_r be the rth prime ($p_1=2$). Write $d_1=p_1$, $d_r=p_r-p_{r-1}$ for $r\geqq 2$. Let D_m be the largest of $d_1, \cdots, d_m$. Then $p_m=d_1+\cdots+d_m\leqq mD_m$, whence $D_m\geqq p_m/m>\log p_m-3/2$ for $p_m\geqq 5$, by a result of Rosser and Schoenfeld (Illinois J. Math., 6 (1962) 64–94; Theorem 2). Therefore, if $n\geqq 1$ and p_m is the smallest prime $\geqq\exp(n+3/2)$, we have $D_m\geqq n$. Now by a theorem of Rohrbach & Weis (J. Reine Angew. Math., 214/215 (1964) 432–440), $p_m\leqq(14/13)\exp(n+3/2)$ for $n\geqq 4$, and hence $p(n-1)\leqq(14/13)\exp(n+3/2)$ for $n\geqq 4$ (the cases $1\leqq n\leqq 3$ being trivial).

Shanks (Math. Comput. 18 (1964), 646–651) conjectured a much stronger result, namely that $\log p(g)\sim\sqrt{g}$ as $g\to+\infty$.

Also solved by Robert Breusch, N. J. Fine, W. D. Fryer, P. Gotiel (Netherlands), M. G, Greening (Australia), D. A. Hejhal, Agnis Kaugars, John B. Kelly, J. G. Mauldon (England), Lieselotte Miller, C. B. A. Peck, Harsh Pittie, R. E. Shafer, Robin Sibson (England), Al Somayajulu, J. Wesley, and the proposer.

When is $n!=a^2+b^2$?

5318 [1965, 796]. *Proposed by Oystein Ore, Yale University*

Prove that there exist only a finite number of integers n for which the factorial $n!$ is the sum of two integral squares.

Solution by Robert Breusch, Amherst College. A prime $p=4m+3$ can divide a^2+b^2 only if $p^2|(a^2+b^2)$. For $n\geqq 7$, there exists a prime $p\equiv 3 \pmod 4$ such that $n/2<p\leqq n$. See Breusch, *Zur Verallgemeinerung des Bertrandschen Postulates dass zwischen x und 2x stets Primzahlen liegen*, Math. Zeitschrift 1932, p. 505. Thus $n!$ cannot be a sum of two squares if $n\geqq 7$. $3!$, $4!$, $5!$ are not sums of two squares because each contains a single factor 3. Thus the only solutions are

$$1! = 0^2 + 1^2, \qquad 2! = 1^2 + 1^2, \qquad 6! = 24^2 + 12^2.$$

Also solved by P. T. Bateman, R. A. Feinman, Harley Flanders (England), Michael Goldberg, Emil Grosswald, J. G. Mauldon (England), K. K. Norton, C. B. A. Peck, Harsh Pittie, Donald Quiring, D. L. Silverman, Al Somayajulu, E. W. Trost (Switzerland), Van de Vyle (Belgium.) and K. S. Williams (England).

Editorial Note. The fact mentioned in the first sentence above is attributed to Fermat and has been rediscovered many times. It may be found in many books—see, e.g., W. J. LeVeque, *Topics in Number Theory*, V. 1, p. 126.

A Class of Space Filling Curves

5321 [1965, 914]. *Proposed by Børge Jessen, University of Copenhagen, Denmark*

Let C denote the class of all pairs (f, g) of continuous functions on the interval $[0, 1]$ on $\mathcal{R}$, such that

$$\int_0^1 [f(t)]^n[g(t)]^m dt = 1/(n+1)(m+1)$$

for all pairs (n, m) of nonnegative integers. Each pair $(f, g)\in C$ determines a continuous curve $(x, y)=(f(t), g(t))$, $t\in[0, 1]$, in $\mathcal{R}^2$. Prove that

(i) The class C is not empty.

(ii) For any pair $(f, g)\in C$, the set $\{(f(t), g(t))\,|\,t\in[0, 1]\}$ is the unit square $[0, 1]\times[0, 1]$ in $\mathcal{R}^2$, i.e., the corresponding curve is contained in the square and contains all its points.

(iii) For any pair $(f, g)\in C$, the functions f and g are nowhere differentiable to the right in $[0, 1($ and nowhere differentiable to the left in$)\ 0, 1]$.

Solution by F. W. Steutel, Enschede, Netherlands. The functions f and g are continuous and therefore bounded on $[0, 1]$. Multiplying the given relations by $x^ny^m/n!m!$ we get (equivalently)

$$(1) \qquad \int_0^1 \exp\{xf(t) + yg(t)\}dt = \frac{e^{x-1}}{x}\,\frac{e^{y-1}}{y}$$

for all x and y. Thus f and g may be interpreted as independent random variables, both uniformly distributed on $[0, 1]$, i.e., if μ denotes Lebesgue measure, we have

$$(2) \qquad \mu(\{t\,|\,f(t)\in\mathfrak{F} \text{ and } g(t)\in\mathcal{G}\}) = \mu(\mathfrak{F}\cap[0, 1])\mu(\mathcal{G}\cap[0, 1]),$$

for all Borel sets $\mathfrak{F}$ and $\mathcal{G}$.

(ii) As f and g are continuous, it follows from (2) that f and g take no values outside $[0, 1]$. On the other hand, the curve $\{f(t), g(t)\}$ is a closed set and therefore the existence of one point of $[0, 1]^2$ not belonging to the curve implies the existence of an open subset of $[0, 1]^2$ not belonging to the curve. This contradicts (2) and proves (ii).

(iii) For an arbitrary point t_0 and any $\epsilon > 0$ there is a $\delta > 0$ such that $|f(t) - f(t_0)| < \epsilon$ if $|t - t_0| < \delta$. By (2) we have

$$\mu(\{t \mid |f(t) - f(t_0)| < \epsilon \text{ and } |g(t) - g(t_0)| < \delta\}) \leqq 4\epsilon\delta.$$

This means that any interval larger than $[t_0, t_0 + 4\epsilon\delta]$ or $[t_0 - 4\epsilon\delta, t_0]$ contains points for which $|g(t) - g(t_0)| \geqq \delta$ and therefore in any such interval

$$\sup \left| \frac{g(t) - g(t_0)}{t - t_0} \right| \geqq \frac{\delta}{4\epsilon\delta} = \frac{1}{4\epsilon}.$$

The quotient $|\{g(t) - g(t_0)\}/(t - t_0)|$ is not bounded in any right or left neighborhood of t_0 (similarly for $f(t)$) which proves (iii).

(i) An example of a pair of functions satisfying (2) is provided by Peano's space filling curve (see e.g., Hobson, *The Theory of Functions of a Real Variable*). Here (2) follows from the fact that all values of t for which the first $2n$ digits in the ternary expansion are constant, $t = 0.\, a_1 a_2 \dots a_{2n} \dots$ are mapped one to one (except for an enumerable set) on the square $f = 0.b_1 b_2 \dots b_n \dots$, $g = 0.c_1 c_2 \dots c_n \dots$.

Also solved by the proposer.

Evaluation of $\phi(d)/e(d)$ (mod 2)

5322 [1965, 914]. *Proposed by John Brillhart, University of San Francisco*

Let $S(n, a) = \sum_{d|n} f(d)$, where $(n, a) = 1$, $f(1) = 1$, and $f(d) = \phi(d)/e(d)$, $d > 1$, ϕ the totient, and $e(d)$ the exponent to which a belongs mod d. Then show that $S(n, 2)$ is odd if $n \equiv \pm 1 \pmod 8$ and is even if $n \equiv \pm 3 \pmod 8$.

Solution by M. G. Greening, University of New South Wales, Australia. It is a simple calculation to show that $\phi(d)/e(d)$ is even, or $f(d) \equiv 0 \pmod 2$, for d a composite number. Therefore $S(n, a) \equiv \sum_{p^t|n} f(p^t)$.

$f(p^t)$ is even if, and only if, $e(p^t)$ is a divisor of $\frac{1}{2}\phi(p^t)$; i.e. if, and only if, $a \equiv \omega^{2k} \pmod{p^t}$, where ω is a primitive root; that is if, and only if, a is a quadratic residue of p. It follows that

$$S(n, a) \equiv \sum_{p^t|n}{}' f(p^t) \equiv f(1) + \sum_{p^\alpha|n}{}' \sum_{t=1}^{\alpha} f(p^t) \qquad \pmod 2,$$

$$(*) \qquad S(n, a) \equiv 1 + \sum_{p^\alpha|n}{}' \alpha \qquad \pmod 2,$$

where $\sum'$ indicates a sum over primes for which a is a quadratic nonresidue.

Now 2 is a quadratic nonresidue for those primes $p \equiv \pm 3 \pmod 8$, and a residue for those primes $p \equiv \pm 1 \pmod 8$. Thus $S(n, 2)-1 \equiv \sum' \alpha \equiv 0 \pmod 2$ if $n \equiv \pm 1 \pmod 8$, and $\equiv 1 \pmod 2$ if $n \equiv \pm 3 \pmod 8$.

Also solved by L. Carlitz, Al Somayajulu, and by the proposer in collaboration with Emma Lehmer.

Brillhart and Lehmer observe that the formula (*) above may be written $S(n, a) \equiv \frac{1}{2}\{(a/n)+1\}$ (mod 2), where (a/n) is the Jacobi symbol. The problem was located in P. Bachmann, *Niedere Zahlentheorie*, problem #609.

Sums and Differences of Two Cubes

5324 [1965, 915]. *Proposed by Alan Sutcliffe, Congleton, Cheshire, England*

Following Hardy's visit to Ramanujan at Putney it is well known that 1729 is the smallest number that is the sum of two cubes in two different ways. (a) Show that 3367 is the difference of two cubes in three ways. (b) What is the smallest number that is the sum of two positive cubes in three different ways?

Solution by John Leech, The University, Glasgow, Scotland.

(a) $$3367 = 15^3 - 2^3 = 16^3 - 9^3 = 34^3 - 33^3,$$

(b) $$87,\ 539,\ 319 = 414^3 + 255^3 = 423^3 + 228^3 = 436^3 + 167^3,$$

and this is the smallest possible as revealed by a computer search. See a tabulation of the five smallest solutions as given in *Some Solutions of Diophantine Equations*, Proc. Cambridge Phil. Soc., 53 (1957) 779.

Also solved by Joseph Arkin, J. A. H. Hunter, and Andrzej Makowski (Poland).

Several solvers quoted 175959000, believed by A. Gérardin to be the smallest number expressible in three ways as the sum of two cubes [L'intermediaire des math. 15 (1908), 182; mentioned in Dickson, *History of the Theory of Numbers*, II, 562]. Makowski gave a reference to C. E. Britton, *On a chain of equations*, Scripta Math., 24 (1959), 179–180, which gives a number expressible as a sum of two cubes in four ways. Numbers (not smallest) are found by Hunter proceeding from the solutions of $x^3-y^3=a^3+b^3$.

Apologies are offered for the unfortunate misprint in the original statement. Most solvers noted that 3367 was the correct number, rather than 3368 as printed.

The Coefficients of $(\sum c_i x^i)^n$

5325 [1965, 913]. *Proposed by D. J. Newman, Yeshiva University and W. E. Weissblum, AVCO Research*

Suppose $0 \leqq c_i \leqq 1$ and $\sum c_i = 2$. Prove that all the coefficients of $(\sum c_i x^i)^n$ are $\leqq \binom{n}{q}$, where $q = [n/2]$.

I. *Solution by L. J. Goldstein, Princeton University.* Proof by induction; the case $n=1$ is clear. Put

$$\left(\sum_{i=0}^{\infty} c_i x^i\right)^n = \sum_{i=0}^{\infty} d_i x^i,$$

and assume that $d_i \leqq \binom{n}{q}$, $q = [n/2]$. Now writing

$$\left(\sum_{i=0}^{\infty} c_i x^i\right)^{n+1} = \sum_{N=0}^{\infty} e_N x^N,$$

we see that

$$e_N = \sum_{i+j=N}^{\infty} d_i c_j \leqq \binom{n}{q} \sum_{i+j=N} c_j \leqq \binom{n}{q} \sum_{j=0}^{\infty} c_j \leqq 2\binom{n}{q} \leqq \binom{n+1}{q'},$$

where $q' = [(n+1)/2]$.

II. *Solution by the proposers.* We prove, in fact, that if $c_i^{(j)}$, $j=1, 2, \cdots, n$, satisfy the above hypotheses, then the coefficients in the power series resulting from

$$\left[\sum c_i^{(1)} x^i\right]\left[\sum c_i^{(2)} x^i\right] \cdots \left[\sum c_i^{(n)} x^i\right]$$

are all $\leqq \binom{n}{q}$. Having fixed the sequences $c_i^{(2)}, \cdots, c_i^{(n)}$ it is clear that one maximizes the largest coefficient in the above product by the choice of $c_i^{(1)}$ equal to 1 for some two values of i. Following this reasoning we see that the maximum coefficient is highest for a product $(1+x^{a_1})(1+x^{a_2}) \cdots (1+x^{a_n})$.

The coefficient of x^N in this product, however, is the number of subsets, S, of $\{1, 2, \cdots, n\}$ for which $\sum_{i \in S} a_i = N$. These subsets have the property that no one contains another and so the result follows from a lemma of Erdös that of any $\binom{n}{q}+1$ subsets of $\{1, 2, \cdots, n\}$ there always exists one which is strictly included in another.

An Arithmetic Inequality

5326 [1965, 915]. *Proposed by C. S. Venkataraman and R. Sivaramakrishnan, Trichur, India*

If $\sigma_k(n)$ denotes the sum of the kth powers of the positive divisors of a positive integer n and $\tau(n)$ is the number of positive divisors of n, prove that

$$\frac{\sigma_k(n)}{\tau(n)} \geqq \sqrt{n^k}.$$

I. *Solution by S. L. Segal, University of Vienna, Austria.* Let $f(n)$ be a positive submultiplicative arithmetic function, $f(nm) \leqq f(n)f(m)$, $n, m \geqq 1$; then

$$\frac{1}{\tau(n)} \sum_{d|n} f(d) \geqq (f(n))^{1/2}.$$

This is a simple consequence of the arithmetic-geometric inequality:

$$(f(n))^{1/2} \leqq \left(\prod_{d|n} f(d)f(n/d)\right)^{1/2\tau(n)} = \left(\prod_{d|n} f(d)\right)^{1/\tau(n)} \leqq \frac{1}{\tau(n)} \sum_{d|n} f(d).$$

The result of the problem now follows upon taking $f(n) = n^k$.

II. *Solution by C. B. A. Peck, Pennsylvania State University.* $a^b(a-1) \geqq a-1$ for all positive a, b with equality only for $a=1$. Setting $a^2 = n/m^2$ with positive m, n, $b=2k-1$, we see that

$$(n/m)^k + m^k \geqq (m^{k-1} + (n/m)^{k-1})\sqrt{n} \qquad \text{for } m^2 \neq n$$

$$m^k = m^{k-1}\sqrt{n} \qquad \text{for } m^2 = n.$$

Specializing k, m, n to positive integers, and adding over all divisors m of n with $1 \leqq m^2 \leqq n$, we get

$$\sigma_k(n) \geqq \sigma_{k-1}(n)\sqrt{n} \geqq \cdots \geqq \sigma_0(n)\sqrt{n^k} = \tau(n)\sqrt{n^k}.$$

The case $k=1$ is problem E 1625 [1963, 891; 1964, 683].

Also solved by A. A. Aheart, P. N. Bajaj, N. J. Fine, J. A. Fridy, L. J. Goldstein, M. G. Greening (Australia), Emil Grosswald, D. A. Hejhal, Erwin Just, Marijo LeVan, Andrzej Makowski (Poland), C. J. Moremo, C. B. A. Peck, Stanton Philipp, Harsh Pittie, Simeon Reich (Israel), Al Somayajulu, Sidney Spital, Preston Stein, M. V. Subbarao, A. M. Vaidya (India), Van de Vyle (Belgium), and the proposers.

Grosswald and Van de Vyle note that the inequality is strict for $n>1$.

Triangle Tesselations of the Plane

5328 [1965, 915]. *Proposed by J. H. Conway, Gonville and Caius College, Cambridge, England*

Show that the plane may be filled with rational-sided triangles in such a way as to use just one of each type. (The triangles should be taken as closed and "filled" means covered in such a way that the overlap has zero area.)

Solution by D. C. Kay, University of Wyoming. Consider the triangle T in E^2 whose vertices are (0, 0), (3, 0), and (0, 4). T has the rationals 3, 4, 5 as sides. We will show how to divide T into two subtriangles having rational sides for every integer $n>5$. If

$$x_n = \frac{16n(n^2+1)}{(n^2-1)^2},$$

since $0<x_n<3$ for each $n>5$ then the point $(x_n, 0)$ may be joined to (0, 4) forming two subtriangles $T'(n)$ and $T''(n)$ having sides 4, x_n, $d_n = (x_n^2+16)^{1/2}$ and 5, $3-x_n$, d_n, respectively, and

$$d_n = \frac{4(n^4+6n^2+1)}{(n^2-1)^2}.$$

Thus d_n is rational for each n, and the triangles $T'(n)$ and $T''(n)$ are nonsimilar for different values of n. Now tesselate the plane by 3, 4, 5 right triangles which may be labelled $T_1, T_2, \cdots, T_n, \cdots$ in some order. The desired tesselation of the plane by distinct triangles with rational sides is effected by $T_1'(6)$, $T_1''(6)$, $T_2'(7)$, $T_2''(7)$, $\cdots$.

Also solved by Michael Goldberg, John B. Kelly, and the proposer.

Kelly's solution involves a demonstration that any triangle with rational sides may be split into two triangles with rational sides in infinitely many different ways.

The interesting question may still be raised: is a tesselation of the required type possible using triangles with integral sides?

Integral Functions of an Integral Function

5329 [1965, 915]. *Proposed by Leopold Flatto, Yeshiva University*

Let $\omega(z)$ be an entire function such that there exists another nonconstant entire function $f(z)$ for which $f(z)=f(\omega(z))$. Show that this implies $\omega(z)=\zeta z+a$, where ζ is a root of unity.

I. *Solution by Fred Gross, Naval Research Lab.* Let $M_f(r)=\max_{|z|=r}|f(z)|$. G. Polya [J. London Math. Soc., I (1926) 12–15, *Integral functions of an integral function.* See also J. E. Littlewood, *Lectures on the Theory of Functions*, pp. 225–226.] shows that if f and g are entire and $g(0)=0$, then $M_{fg}(r)=M_f(cM_g(r/2))$ for an appropriate constant c, $0<c<1$. It follows that if $f(z)=f(\omega(z))$, then $\omega(z)$ must be linear, say $\omega(z)=\zeta z+a$. Write

$$F(\zeta z)=f\left(\zeta\left(z+\frac{a}{1-\zeta}\right)+a\right)=f\left(z+\frac{a}{1-\zeta}\right)=F(z).$$

The Taylor series of $F(z)$ shows that ζ is a root of unity.

II. *Solution by R. Goldstein, The Northern Polytechnic, London, England.* The problem can be generalized and is valid if f is a nonconstant meromorphic function. A result of Rosenbloom [*The fixed points of entire functions*, Medd. Lunds Univ. Mat. Sem. Tome Suppl. M. Rieve, (1952) 186–192] says that either $\omega(z)\equiv z+a$, $a\neq 0$, or $\omega^{(2)}(z)\equiv\omega(\omega(z))$ has a fixed point, for any entire function $\omega(z)$. So we may suppose that there exists a complex number b such that $\omega^{(2)}(b)=b$. Let

$$(1)\ h(z)\equiv\omega(z+b)-b,\qquad (2)\ F(z)\equiv f(z+b).$$

Then, by hypothesis

$$F(z)=f(z+b)=f(\omega^{(2)}(z+b))=f(h(z)+b)=F(h(z)).$$

Thus $z=0$ is a fixed point of the entire function $h(z)$ and there is a nonconstant meromorphic function $F(z)$ such that

$$F(z)=F(h(z))=\cdots=F(h^{(n)}(z))=\cdots, \tag{3}$$

for all z, and for all integers n, where $h^{(1)}(z)=h(z)$, $h^{(n+1)}(z)=h(h^{(n)}(z))$, $n\geqq 1$. Without loss of generality we may suppose that $F(z)$ is regular in the neighborhood of the origin and that $F(0)=0$. Let

$$F(z)=b_pz^p+b_{p+1}z^{p+1}+\cdots,\qquad b_p\neq 0,\ p\geqq 1 \tag{4}$$

in some neighborhood of the origin, and let

(5) $$h(z) = a_k z^k + a_{k+1} z^{k+1} + \cdots = a_k z^k[1 + O(z)], \qquad a_k \neq 0,\ k \geqq 1.$$

Then

(6) $$\begin{aligned} F(h(z)) &= b_p a_k^p z^{kp}[1 + O(z)]^p + b_{p+1} a_k^{p+1} z^{kp+k}[1 + O(z)]^{p+1} \cdots \\ &= b_p a_k^p z^{kp} + O(z^{kp+1}). \end{aligned}$$

So, comparing (4) and (6), using (3), gives $k=1$ and

(7) $$a_1^p = 1.$$

Thus $h(z) = a_1 z + O(z^2)$, $a_1 \neq 0$. So

$$h^{(2)}(z) = a_1(a_1 z + O(z^2)) + O(z^2) = a_1^2 z + O(z^2),$$

and, in general

(8) $$h^{(p)}(z) = a_1^p z + O(z^2) = z + O(z^2) = z + c_k z^k + \cdots, \qquad k \geqq 2$$

say. Then

(9) $$\begin{aligned} F(h^{(p)}(z)) = b_p z^p[1 + c_k z^{k-1} + \cdots]^p &+ b_{p+1} z^{p+1}[1 + c_k z^{k-1} + \cdots]^{p+1} \\ &+ b_{p+k-1} z^{p+k-1}[1 + c_k z^{k-1} + \cdots]^{p+k-1} + \cdots. \end{aligned}$$

Comparing the coefficients of z^{p+k-1} in (4) and (9) gives

$$b_p p c_k + b_{p+k-1} = b_{p+k-1}.$$

Hence $b_p p c_k = 0$, whence $c_k = 0$ ($b_p \neq 0$, $p \neq 0$).

Thus $h^{(p)}(z) = z$ in the neighborhood of the origin, and so also in the whole plane. By the last equation $h^{(p)}(z)$ is univalent. Therefore $h(z)$ must be univalent too. But $h(z)$ is an entire function at the same time, so $h(z)$ must reduce to a linear function of z. As $h(0) = 0$ we must have

(10) $$h(z) = a_1 z, \qquad a^p = 1.$$

Then, by (1)

(11) $$\omega^{(2)}(z) = h(z - b) + b = a_1(z - b) + b = a_1 z + b(1 - a_1).$$

Thus $\omega^{(2)}(z)$ is univalent, so $\omega(z)$ is univalent too. As $\omega(z)$ is also entire, we must have $\omega(z) = \zeta z + a$. So

(12) $$\omega^{(2)}(z) = \zeta^2 z + \zeta a + a.$$

Comparing (11) and (12) we have $\zeta^2 = a_1$, whence $\zeta^{2p} = 1$. Thus $\omega(z) = \zeta z + a$ with $\zeta^{2p} = 1$ as required.

It may happen that $\omega(z)$ itself has already a fixed point $z = b$. Then, by putting $h(z) = \omega(z + b) - b$ we may obtain $\omega(z) = \zeta z + a$ with $\zeta^p = 1$.

Also solved by I. N. Baker (England).

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: KENNETH O. MAY, University of Toronto and E. P. VANCE, Oberlin College

Materials intended for review should be sent to Kenneth O. May, American Mathematical Monthly, Department of Mathematics, University of Toronto, Toronto 5, Ont., Canada.

The Riemann Hypothesis and Hilbert's Tenth Problem. By S. Chowla. Gordon and Breach, New York, London, Paris, 1965. 119 pp. $6.50.

Put

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

where $\chi(n)$ denotes a character (mod k). The Dirichlet L-function $L(s, \chi)$ can be extended over the entire s-plane and is an entire function for χ a nonprincipal character. The extended Riemann hypothesis asserts that the so-called nontrivial zeros of $L(s, \chi)$ lie on the line $R(s) = \frac{1}{2}$. Hilbert's tenth problem asks for a finite algorithm to solve a diophantine equation $f(x, y) = c$, where f is a polynomial in x and y with, say, rational integral coefficients. The subjects discussed in the present book are at least loosely related to one or the other of these topics. In a little over a hundred pages, Professor Chowla discusses an amazing variety of interesting questions in number theory. The first four chapters are concerned with character sums, the sign of the Gauss sum, the least positive quadratic nonresidue of a large prime and the formula of Jacobsthal:

$$a = \frac{1}{2} \sum_{x=1}^{p} \left(\frac{x}{p}\right)\left(\frac{x^2 - 1}{p}\right),$$

where $p = a^2 + b^2$, a odd. Chapter 5 contains a proof of the congruence

$$\pm 2a \equiv \binom{2m}{m} \pmod{p} \qquad (p = 4m + 1).$$

Chapters 6 and 7 are concerned with diophantine equations. Among other things there is a proof of the Mordell-Weil theorem that all rational points on the curve

$$y^2 = (x - \theta_1)(x - \theta_2)(x - \theta_3),$$

where the θ_i are rational integers, are obtainable from a finite number by the tangent chord process. Chapter 8 contains a list of sixty-nine problems and theorems of varying degrees of difficulty. This chapter will no doubt be of special interest to the reader.

The final chapter is concerned with cyclotomy; in particular it contains a

proof of the theorem that if p is a prime $\equiv 1 \pmod 3$, then 2 is a cubic residue $(\bmod\ p)$ if and only if $p=x^2+27y^2$.

This little book is very readable and should prove very valuable to the student of number theory.

L. CARLITZ, Duke University

Addition Theorems: The Addition Theorems of Group Theory and Number Theory. By Henry B. Mann. Interscience, New York, 1965. 114 pp. \$8.75.

In 1942 Professor Mann proved the following theorem that had been long outstanding. Let A, B, C be sets of positive integers. Let $A(n)$ denote the number of integers in A that do not exceed n; $B(n)$, $C(n)$ are defined in a similar way. Let

$$C^0 = \{a + b \mid a \in A^0, b \in B^0\},$$

where A^0, B^0 consist of the numbers of A and B, respectively, and the number 0. Then

$$\frac{C(n)}{n} = 1 \quad \text{or} \quad \geqq \min \frac{A(m) + B(m)}{m} \qquad (1 \leqq m \leqq n,\ m \notin C).$$

Extensions of this result were obtained by the author, F. J. Dyson, M. Kneser, A. M. Macbeath and others. The first four chapters of this book contain a unified account of these results. Chapter 5 is concerned with theorems about bases of the integers proved by Erdös, Kasch and Landau. Chapters 6, 7, 8 are devoted to difference sets. The final chapter on decomposition theorems treats the following question. Let

$$C = A + B = \{a + b \mid a \in A, b \in B\}.$$

If C and A are given, under what conditions will a set B exist that satisfies this equation?

This little book is not easy to read but will be of great value both to the student and to the specialist.

L. CARLITZ, Duke University

Mathematical Methods in Engineering and Physics. Special Functions and Boundary-Value Problems. By David E. Johnson and Johnny R. Johnson. Ronald Press, New York, 1965. viii+271 pp. \$9.50.

This book is designed for courses in applied mathematics in which emphasis is placed on acquainting the student with the characteristics of the more important special functions used in solving partial differential equations encountered in engineering and physics. Attention is directed primarily to boundary-value problems, with emphasis on the separation-of-variables method. A brief introduction to transform methods is included. Applications to boundary-value problems in engineering and physics are discussed as they arise in the examinations of the functions themselves.

A good working knowledge of elementary differential equations is a necessary prerequisite for the proper study of this book. Such topics as are needed from advanced calculus are considered as they arise. The materials studied are considered generally to be on an advanced level though they are presented in a form suitable for students with less extensive backgrounds. The book may be used for one or two-semester courses at the graduate or senior level.

The book begins with a study of orthogonal functions and a statement of the Sturm-Liouville problem. Self-contained chapters on Fourier series and the series solution of differential equations are included. Next are chapters on the Legendre, Gamma and Bessel Functions; Hermite, Laguerre, and Chebyshev Polynomials; Mathieu and other special functions. A chapter on concepts underlying boundary-value problems precedes one on Partial Differential Equations of Mathematical Physics, in which the common equations of mathematical physics are listed (i.e., Helmholtz, wave, diffusion, Laplace's, vibrating string and membrane equations). Here, as elsewhere, applications of special functions to the boundary-value problems are included. The book closes with introductory chapters on Laplace, Fourier, and Sturm-Liouville Transforms.

This reviewer questioned whether certain unifying concepts might have been introduced early from which many of the characteristics of special functions may be obtained. This possibility was considered by the authors who state in the last section that a unifying approach is conceivable and "would have the merit of making the work more compact, but would perhaps give the reader less insight and certainly much less experience in dealing with the special functions."

Although the authors recognize the theories necessary to justify techniques used, they make no attempt to go deeply into such matters. The inclusion of more detailed examples in the chapters on transforms would have been helpful to the student. The book is directed to the development of mathematical techniques rather than to related physical principles. The student may look elsewhere to acquire any needed knowledge of the physical sciences (e.g., the derivations given for the equations for vibrating membranes, and diffusion, are not very convincing, though the student familiar with the physical problems involved will have no difficulties).

The book is carefully written, of pleasing appearance, and should prove to be a useful text. The student who masters it should have developed a good working knowledge of special functions encountered in physical sciences.

R. S. BURINGTON, Arlington, Virginia

Applications of Characteristic Functions. By E. Lukacs and R. G. Laha. Griffin's Statistical Monographs and Courses No. 14. Hafner, New York, 1964. 202 pp. $7.25.

This monograph might be considered as a sequel to Lukacs "Characteristic Functions" published in 1960 as No. 5 of the Griffin Statistical Monograph series, and reviewed by this writer in the MONTHLY, 69 (1962) 820–821. Where-

as the earlier volume emphasized the theory of characteristic functions, the present volume is primarily concerned with applications to problems of theoretical statistics. The first chapter constitutes an introduction, while subsequent chapters are titled as follows: (2) "Important multivariate distributions," (3) "The distribution problem of statistics," (4) "Independent statistics in a given population," (5) "Characterization of populations by the independence of statistics," (6) "Regression problems," (7) "Some problems in the theory of stochastic processes," (8) "Identically distributed statistics," (9) "Characterization of populations by the distribution of a statistic," and (10) "Stability theorems." There are six appendices bearing titles as follows: (A) "Some basic concepts of probability theory," (B) "Fourier transforms of probability measures in R_n," (C) "Some results concerning determinants and matrices," (D) "Proof of a lemma used in chapter 3," (E) "Proof of lemmas used in chapter 5," and (F) "Stochastic convergence."

Although detailed proofs are given for most of the theorems considered, occasionally proofs are omitted and reference is made to the original papers. In particular, only a brief indication is given of the structure of the proofs for some of Yu. V. Linnik's work concerning identically distributed statistics.

One of the noteworthy contributions of this and the earlier monograph is that these works make available many important results that heretofore have been published only in Russian and in other foreign journals. The scholarship exhibited in this work is commendable and the authors deserve the thanks of the entire statistical community for making it available.

A. C. COHEN, JR., University of Georgia

Introductory Probability and Statistical Applications. By P. L. Meyer. Addison-Wesley, Reading, Mass., 1965. x+339 pp. $8.75.

This carefully written text, presupposing one year of calculus, is directed towards engineering students. It covers the usual topics of non-measure-theoretic probability as well as containing brief excursions into hypothesis testing, statistical estimation, and reliability. The latter makes a valuable addition to engineering courses, brief though it is, but the chapters on statistical ideas are too condensed to be useful. The author states that the book is intended for a one-semester course; my experience has been that there is too much material here for that time. One can "get through" the material, but one would do well to "teach" the first nine chapters and chapter 12, a total of 210 pp., with selected illustrations from the chapters on moment generating functions and reliability.

The book contains a great many well thought out and imaginative examples and problems, and no obvious misprints. It is both more readable and better organized, and hence more teachable, than any other text covering the same material (probability—not statistics) that comes to mind. A question that will not down, however, is when a course covering this material might be appropriate, especially for engineers. For those with only one semester to spare, a com-

bined course with more statistics content would seem to be indicated. For those taking a full year course, there are better statistical texts. Nevertheless, if you *have* to teach a short course on probability to engineers, this will be an eminently satisfactory book.

ROBERT KOZELKA, Williams College

Boundary Value Problems. By A. G. Mackie. Hafner, New York, 1965. 252 pp. $5.50.

The aim of this book is to give a general account of how certain mathematical techniques may be used to solve important classes of boundary value problems in ordinary and partial differential equations. Emphasis is placed on the use of Green's functions and integral transform methods. The approach of the author has been essentially heuristic without insistence on full rigor. For instance, the ideas of delta functions are used without development of a rigorous theory.

In Chapters 2 and 3 the author concerns himself with solutions of ordinary differential equations satisfying one and two point boundary conditions. Transform methods, and techniques involving delta functions, Green's functions, and eigenfunction expansions are introduced and applied to a wide variety of problems. Chapters 3–6 deal with solutions of various boundary value problems for partial differential equations. In Chapter 3 simple equations of various types are introduced and methods of separation of variables and integral transforms are discussed. Chapter 4 is devoted to Fourier integrals and Fourier, Mellin, and Hankel transforms. The roles of the Green's function, the method of images, and conformal mapping in the treatment of various classes of problems is discussed in Chapter 5. Finally, Chapter 6 is concerned with Riemann's method for solving hyperbolic problems and the connection between Riemann's method and the method of Green's functions.

The book is clearly written and well within the grasp of the well-trained senior undergraduate mathematician, physicist or engineer. It should prove useful as a text or reference for a senior course in mathematical methods.

L. E. PAYNE, Cornell University

A Vector Space Approach to Geometry. By Melvin Hausner. Prentice-Hall, Englewood Cliffs, N. J., 1965. 397 pp. $12.00.

Finding a title for a mathematical book is not easy these days, and since many books are adopted by departmental chairmen at sight, as it were, after a mere scrutiny of title, name of author and table of contents, it behooves an author to choose a title which is not too out of phase with the prevailing fashion in mathematical texts. The title to this very attractive book does not describe the contents. In fact, the first chapter deals with physical motivations in geometry, and the preface says: "although much algebra is used and developed here, this book is basically a geometry text, and no serious attempt has been made to give a totally axiomatic account of vector spaces."

The book is based on one year in-service courses for high school teachers of

mathematics given by the author, and explores the effect that geometry and linear algebra have on each other. It is carefully written and thought out, and can be recommended for similar courses where the instructor wishes to introduce his students to some of the delights of mathematics, rather than to crush them by asking them to gaze up at a towering superstructure, which they will have to climb in the course of a year without the benefit of elevators.

The chapter headings are: The center of mass, Vector algebra, Vector spaces and subspaces, Length and angle, Miscellaneous applications, Area and volume, Further generalizations, Matrices and linear transformations, Area and metric considerations, The algebra of matrices, Groups.

DANIEL PEDOE, University of Minnesota

Linear Algebra, Volume III, Part 1. By V. I. Smirnov, translated by D. E. Brown and edited by I. N. Sneddon. Addison-Wesley, Reading, Mass., 1964. ix+324 pp. $9.00.

Appearing as the first part of Volume III of a five volume set by the author entitled *A Course of Higher Mathematics*, the purpose is to present the fundamentals of linear algebra and the theory of groups that are "most frequently used in theoretical physics." Noting the sectional topics mentioned below, an idea of the breadth of the treatment is obtained. The choice of topics should satisfy the requirements for a background in matrices and their applications for physicists as well as being topics that should supplement any first course in matrix theory.

The book is quite readable, there are many fine examples, and no exercises. The approach used is antiquated when measured against today's trends of emphasis and it is pronounced enough to act as a drawback when considering it for an introductory undergraduate course, in particular one directed toward mathematics majors. Containing a wealth of topics not normally treated in a first book of matrix algebra, a revision would make it a superior text.

In Chapter I determinants are introduced in the classical manner followed by the solutions of systems of equations. Flavor is added by including sections on Gram's determinant and Hadamard's inequality, systems of linear differential equations with constant coefficients, functional determinants (Jacobians), as well as an outlined proof of the implicit function theorem.

Chapter II is concerned with linear transformations and quadratic forms. Under current usage, the term linear substitutions would be more appropriate since the entire chapter is developed around this notion. The treatment is enhanced though by having sections on Buniakowski's inequality, small vibrations, functions of matrices, infinite-dimensional spaces, the connection between functional and Hilbert space, and linear functional operators.

The basic theory of groups and linear representations of groups appear in Chapter III. Among the sections of this chapter one finds continuous groups, representations of the Lorentz group, and integration over groups.

HOMER BECHTELL, University of New Hampshire

Principles of Computation. By Peter Calingaert. Addison-Wesley, Reading, Mass., 1965. 200 pp. $7.75.

This book is an introduction to methods commonly used in automatic computation. It begins with a historical introduction on the development of computer machines which is interesting, if not altogether accurate (e.g., the IBM Card Programmed Computer preceded the Ferranti Mark I and the Univac as a commercially available electronic computer).

Dr. Calingaert clearly defines much of the computer terminology, such as an interpreter, compiler, assembler, loop, object program, compiling program, precision, accuracy, error, roundoff error, truncation. He illustrates the development of computer logic by examples and emphasizes the importance of efficiency in programming by presenting alternate ways of programming a simple sort routine. In an attempt to be explicit, the author sometimes makes simple things complex; e.g., the discussion of iteration control seems to be much too detailed.

The book includes a good description of number bases and the transformation from one basis to another. It also touches upon the highlights of probability and statistics, including frequency distributions (discrete and continuous), kinds of errors, correlations, and linear regressions. A section on analog computation includes methods of constructing nomograms, mechanical analog devices, and electronic analog computer programs.

Finally, a section on numerical approximation is included. It is a good introduction, but does not go into depth on any method. For example, only one short paragraph and one short example are given on the predictor-corrector method for a numerical solution of a differential equation. No treatment at all is given to systems of differential equations.

J. C. ROGERS, Omnimetrics

Information and Information Stability of Random Variables and Processes. By M. S. Pinsker. Holden-Day, San Francisco, 1964. 256 pp. $10.95.

This is a research monograph accounting for recent (up to 1960, the date of the Russian edition) work in information theory. Most of this work has only appeared in Russian periodicals, to a large extent in articles of R. L. Dobrushin and the author. The book consists of three parts. Part I, Information and information stability of random variables, is devoted to the basic properties of information in far greater generality than in earlier expositions, which were usually restricted to discrete time and space parameter. The key results here are theorems of Dobrushin and of Gelfand, Yaglom, and Perez, concerning entropy and entropy density in general probability spaces. The proofs are omitted in the text (being available in the literature) but A. Feinstein has provided proofs as well as comments on these and other selected results in a sequence of valuable appendices to the various chapters. Part II, entitled Information rate and information stability of random variables, lays the foundations for Part III which is the core of the book. Its title is "Information, information rate, and informa-

tion stability of Gaussian random variables and processes." Here the information theory of Gaussian processes is related to their spectral theory. This part contains new methods permitting explicit computations of entropy, entropy rate, moments of the entropy density, as well as criteria for entropy stability and for the convergence of the distribution of the entropy density to a normal distribution.

The reader is assumed to be familiar with measure theory and with stationary stochastic processes. The exposition is lucid, and compatible with these prerequisites, but the book is hard to read because of a two-fold lack of motivation. The non-specialist will have to take on faith the statement that the results of Part III can be used to solve problems of mathematical statistics concerning random processes, in particular the problem of receiving a signal in the presence of noise. No attempt is made to describe such problems. On the other hand the role of information in related areas of mathematics, such as ergodic theory, could also motivate the contents of the book, but even the appendices provide only hints in this direction. Indeed the subject matter of the book provides a natural setting for many interesting related results such as Linnik's information-theoretical proof of the central limit theorem, and the dichotomy theorem for Gaussian measures of Feldman and Hajek.

The translation is of uniformly high quality, and the appendices mentioned above render the book more valuable than the previous Russian and German editions.

FRANK SPITZER, Cornell University

Mathematics: Its Content, Methods, and Meaning, volumes I, II, III. Edited by A. D. Aleksandrov, A. N. Kolmogorov and M. A. Lavrentev. Translated by S. H. Gould. The M.I.T. Press, Cambridge, Mass., 1964. 359 pp. 377 pp. pp. 356 pp. Set $30.00.

The twenty chapters in these three volumes were written by eighteen Russian mathematicians, many of whom have international reputations. The object is to convey the flavor and some of the content of a large portion of mathematics to an audience consisting of students and teachers as well as engineers, scientists, and intelligent laymen. The preface states that only secondary-school mathematics is presupposed but that for most of the book a working knowledge of calculus, presented in Chapter II, is necessary.

While Chapter II, which gives an uninspired presentation of the usual formulas of elementary calculus, is intended to serve as reference material for other chapters, the book as a whole is in no sense a reference work; there is no index and sometimes the content of a well-known theorem is established, but the theorem is not named or even formally stated. There are no exercises and few purely mathematical examples, but every effort is made to connect the mathematics with its applications and/or sources in physics and engineering. The chapters cover topics ranging from elementary calculus to partial differential

equations, functional analysis, electronic computing machines, and groups and other algebraic structures.

We assume that the book is intended for readers who will be reading near the boundary of their comprehension. Ideally then the exposition should be brilliant and clear and unmarked by flaws. Some chapters come close to attaining this ideal, but the book as a whole does not. To begin with, there are many misprints throughout.

Secondly there are some statements so oversimplified that they are not quite true; and there are statements of questionable validity. We cite a few examples. The chapter on complex variable asserts that every elementary function is analytic at each point of its domain. The chapter on prime numbers asserts that the infinitude of the sequence of natural numbers reflects the infiniteness of the material world in time and space. The chapter on probability asserts that if one computes the frequency of male births during a conjunction of Mars and Jupiter the result would not be different than at other times. This assertion is then used as an illustration of how science refutes the imagined relations of the astrologers.

In some chapters there are too many instances of careless writing and editing. Other chapters are of doubtful value because they cover too much material too thinly. In this category, we place Chapters IX (Functions of a Complex Variable), XV (Theory of Functions of a Real Variable), XVI (Linear Algebra), XX (Groups and Other Algebraic Systems).

On the other hand, some chapters are examples of excellent expository writing. These include the chapters on analytic geometry, the theory of algebraic equations, curves and surfaces, calculus of variations, non-Euclidean geometry and functional analysis. In particular, Chapter III (Analytic Geometry) provides excellent collateral reading of a concrete sort for a student first studying the abstract theory of linear spaces.

Chapter VII on curves and surfaces and Chapter VIII on the calculus of variations are conspicuously successful. From these a reader with a year of calculus should be able to get a good feeling for curvature, torsion, Gaussian curvature and should also be able to learn what it means to minimize a functional, while anyone who understands partial derivatives should be able to follow the derivation of Euler's necessary condition.

Chapter XIX (Functional Analysis) is an outstanding one, although the reader needs a somewhat greater background than a one-year calculus course.

It is of course, not surprising that a book with nineteen authors should be uneven in quality. Unfortunately, this unevenness makes the present work only partially successful. We believe that the problem of communicating mathematics to the intelligent layman is an important and difficult task. The best parts of these three volumes can well serve as models for others who wish to contribute to this task, while the worst parts can serve to warn them how easy it is to miss the mark.

H. E. Chrestenson, B. Hunt, J. B. Roberts, Reed College

Introductory Computer Methods and Numerical Analysis. By Ralph H. Pennington. Macmillan, New York, 1965. 464 pp. $9.00.

This is a book directed toward the student engineer and scientist with a background in mathematics through integral calculus but with no familiarity of the digital computer. The numerical analysis portion of the book is computer-oriented with special emphasis given to possible errors that may develop in the course of calculation. Examples are numerous and well employed. Each chapter includes a set of exercises, most of which are computational in nature.

After the initial chapter on number systems, a brief discussion of a hypothetical computer is given. The reader is then led smoothly through the process of machine language programming to basic FORTRAN-type programming and flow charts. Several chapters are devoted to the important problem of accuracy in computer calculations. The remaining chapters cover the standard introductory numerical methods for digital computation. Included are chapters on evaluation of functions, quadrature, solution of algebraic and transcendental equations, solution of polynomial equations, systems of linear equations, matrices, curve fitting, interpolation and differentiation, and ordinary differential equations. Flow charts and FORTRAN programs are included whenever possible.

In general, the author's exposition is very clear. Due to the amount of material covered, however, some numerical methods are discussed only briefly and are presented as "cook book" procedures. A bibliography is given but, unfortunately, no references are cited in the text. One not so obvious misprint was noted. Page 327 should read " . . . all the eigenvalues . . . ," not " . . . all the elements . . . ."

This book should be useful as a text for either a laboratory or classroom course in introductory digital computing and numerical methods.

L. A. HAGEMAN, Westinghouse Electric Corp.

Integration, Measure and Probability. By H. R. Pitt. Hafner, New York, 1963. 106 pp. $4.00.

This small book seems to be directed toward the graduate student who has already had solid work in linear algebra and matrix theory, and desirably some work in probability and/or mathematical statistics. Its principal objective appears to be the consideration of a variety of theorems concerning the limiting forms of distributions. In Part One (45 pages) the author provides a compact, self-contained introduction to completely additive set functions over a σ-ring of sets, and of measure and integration. The Radon-Nikodym theorem concerning the decomposition of a completely additive set function into its continuous and singular components is stated and proved.

Part Two begins, in Chapter 4, with classical probability theory, based on the set-theoretical approach due to Kolmogoroff. The standard distributions, discrete and continuous, one- and multi-dimensional, are treated, including

random vectors with a general normal distribution. The author proves, for example, the independence of the sample mean and variance for random samples from a normal distribution.

Chapter 5 first deals with the so-called central limit problem, that is, with generalizations and extensions of the central limit theorem. Several theorems are derived which give necessary and sufficient conditions in order that the limiting form of a distribution, as $n \to \infty$, may be of a specified type: normal, Poisson, singular, etc. The chief tools are the characteristic function and the K-L function, due to Kolmogoroff, Khintchine and Lévy. A final section of 13 pages gives a *résumé* of results concerning random functions and sequences, stochastic processes and time series, largely based on the work of such men as Wiener and Doob.

H. W. ALEXANDER, University College, Nairobi, Kenya

Linear Algebra. By R. A. Beaumont. Harcourt, Brace & World, 1965. 216 pp. \$3.50.

The text, one of a series, was designed to be used following a course in Calculus and preceding one in multivariate calculus. I found it to be very interesting. Topics are presented in an orderly fashion, figures are neatly drawn, and the page layout is excellent.

The book is divided into thirty-three sections. Since each section covers about six pages, it appears suitable for a section-a-lesson coverage in a one-semester course meeting three times per week. There are an adequate number of examples and exercises. The exercises involve both proofs and routine computations. The former category includes some proofs of theorems required in the text. Answers and hints are given for many of the exercises.

Chapter headings are: Vector Methods in Geometry; Real Vector Spaces; Systems of Linear Equations; Linear Transformations and Matrices; Equivalence of Matrices; Determinants; Similarity of Matrices; and Quadratic Forms. The notion of real vector spaces is introduced early, and the terminology is used throughout the book.

I believe the coverage is adequate for an introductory course in Determinants and Matrices, and it would certainly be a most helpful prelude to a more abstract course in Modern Algebra.

A. B. FARNELL, General Dynamics/Convair

Ordinary Differential Equations. By Philip Hartman. Wiley, New York, 1963. xiv+612 pp. \$20.00.

The author has selected those aspects of differential equations with which he has been most intimately associated over the last twenty years, added appropriate introductory material, very good exercises, historical notes, and an extensive bibliography to produce an excellent treatise. The chapters on differential inequalities, total and partial differential equations, invariant manifolds and linearizations, perturbed linear systems, applications of fixed point theorems,

dichotomies, and monotony contain much material previously available only in research papers, many of which are by Hartman or Wintner. Although theoretically an advanced undergraduate or beginning graduate student could read through the book, the extensive and very thorough analysis of most of the proofs make the book more appropriate as a reference for a research mathematician. A stimulating graduate course might well be based, however, on selections from Hartman's book and the chapters on stability from Lefschetz's book *Differential Equations: Geometric Theory*, Interscience, 1963.

COURTNEY COLEMAN, Harvey Mudd College

Homographies, Quaternions, and Rotations. By Patrick Du Val. Oxford Mathematical Monographs, Oxford at the Clarendon Press, 1964. xiv+116 pp. 35s.

By means of stereographic projection, a projectivity of the real unit sphere onto itself corresponds to a transformation of the conformal plane which maps the set of all the circles and straight lines onto itself. If the projectivity is orientation preserving, the transformation is a *homography* [=Moebius transformation=linear fractional transformation], otherwise the projectivity determines an *antihomography*. Chapter I collects some information on the groups $\mathfrak{H}$ of all the homographies and $\mathfrak{H}^*$ of all the homographies and anti-homographies and on certain subgroups. Chapter II determines the finite subgroups of $\mathfrak{H}$ and $\mathfrak{H}^*$ and their structure. Quaternion representations are used in Chapter III for the enumeration of the finite subgroups Γ of the four-dimensional real orthogonal group. An elementary argument yields all the possible regular polytopes and tessellations in n-space. The four-dimensional case and the 3-sphere modulo some Γ are then studied in detail (Chapter IV). Chapter V: If $\mathfrak{G}$ is a finite subgroup of $\mathfrak{H}$ and t is any point in the conformal plane, the *group set* $\mathfrak{G}t$ determines an invariant form. For given $\mathfrak{G}$, any three of these forms are linearly dependent. Thus every group set satisfies an algebraic equation $f(u, v)+\mu g(u, v)=0$ in homogeneous coordinates u, v [$\mu=\text{const}$] and the group sets of $\mathfrak{G}$ form a rational involution which is discussed for each $\mathfrak{G}$. $\mathfrak{G}$ is the 2:1 or 1:1 homeomorphic image of a finite centro-affine group $\mathfrak{G}_A$. The group sets of $\mathfrak{G}_A$ are in 1-1 correspondence with the points of an algebraic surface in affine 3-space with a singularity at the origin.

PETER SCHERK, University of Toronto

Numerical Methods: 1. Iteration, Programming and Algebraic Equations. By Ben Noble. Wiley, New York, 1965. xii+156 pp. \$2.75.

Numerical Methods: 2. Differences, Integration and Differential Equations. By Ben Noble. Wiley, New York, 1965. viii+216 pp. \$3.00.

These little books provide an informal, accurate introduction to numerical methods. The areas covered in each volume are indicated in the titles. Tedious details and lengthy derivations are circumvented through footnoted references to elementary sources.

The main concern is to expose basic ideas, and the topics are chosen accordingly. Among the omissions are the $Q-D$ algorithm and Romberg's method, and there is no bibliography. Yet the shortcomings are outweighed by the author's clarity, easy informality, and skill in selecting illustrations.

NATHANIEL MACON, Institute for Defense Analyses

The Haar Integral. By Leopoldo Nachbin. Van Nostrand, Princeton, N. J., 1965. xii+156 pp. $6.50.

Despite the number of existing books about topological groups there is definitely a place for this one. Addressed to graduate students "having the mathematical maturity normally expected from them," and presupposing only a rudimentary knowledge of general topology, algebra, and elementary integration, its aim is to reach the existence and uniqueness theorems for the Haar integral by the shortest possible path. It is hard to imagine how this aim could be more successfully achieved. The exposition moves smoothly and self-sufficiently through the theory of the Radon integral in a locally compact space, then proceeds to locally compact groups. The main existence and uniqueness theorem is stated and assumed, while properties of the Haar integral are derived from it and abundantly illustrated with examples from the classical groups. Then two proofs are presented, following A. Weil and H. Cartan, respectively. A short final chapter on locally compact homogeneous spaces reveals the applications of the Haar integral to integral geometry.

The exposition is strictly according to Bourbaki, and the book provides an excellent introduction to this theory of integration. Thus measure plays only a subordinate role; there is no need to distinguish between Baire and Borel measures, and questions of regularity or of the possibility of invariant extensions do not even arise. There are no lists of problems or exercises. One might criticize the book on the ground that it makes the subject appear almost too finished, but within the limits the author has set himself he has done a superb job. Technically there is almost nothing to criticize, unless it be to point out that the words "less than" sometimes mean "less than or equal to" (e.g. in Propositions 23 and 24, Chap. 1), and that in Lemma 1, p. 32, the condition should read "it is necessary and sufficient that f be finite valued and $\mu^*(f) < +\infty$." But these are very minor points. Here is a book that can be highly recommended as a model of smooth exposition, a pleasure to read.

J. C. OXTOBY, Bryn Mawr College

BRIEF MENTION

Six-Figure Logarithmic Trigonometrical Functions of Angles in Degrees and Minutes, 5th rev. ed. Practical Tables Series No. 1. By C. Attwood. Pergamon Press, Long Island City, N.Y. 1965. 68 pp. $1.50.

Six-Figure Trigonometrical Functions of Angles in Hundredths of a Degree. Practical Tables Series No. 2. By C. Attwood. Pergamon Press, Long Island City, N.Y. 1965. 103 pp. $2.45.

Six-Figure Logarithmic Trigonometrical Functions of Angles in Degrees and Minutes, 5th ed. Practical Tables Series No. 3. By C. Attwood. Pergamon Press, Long Island City, N.Y. 1965. 75 pp. $1.50.

Six-Figure Logarithmic Trigonometrical Functions of Angles in Hundredths of a Degree. Practical Tables Series No. 4. By C. Attwood. Pergamon Press, Long Island City, N.Y. 1965. 100 pp. $1.50.

Six-Figure Logarithms, Cologarithms and Antilogarithms. Practical Tables Series No. 5. By C. Attwood. Pergamon Press, Long Island City, N.Y., 1965. 125 pp. $2.45.

Studies in Real and Complex Analysis, vol. 3 of MAA Studies in Mathematics. Edited by I. I. Hirschman, Jr., Prentice Hall, Englewood Cliffs, N.J. 1965. 213 pp. $4.00.

The articles in this volume are in general directed to advanced undergraduate and first year graduate students. They are, in addition to an introduction by the editor, "Several Complex Variables" by H. J. Bremermann, "Nonlinear Mappings Between Banach Spaces" by L. M. Graves, "What is a Semi-Group?" by Einar Hille, "The Laplace Transform, The Stieltjes Transform and their Generalizations" by I. I. Hirschman, Jr. and D. V. Widder, "A Brief Introduction to the Lebesgue-Stieltjes Integral" by H. H. Schaefer, "Harmonic Analysis" by Guido Weiss, and "Toeplitz Matrices" by Harold Widom.

Introduction to the Foundations of Mathematics, 2nd ed. by Raymond L. Wilder. Wiley, New York, 1965. xvi+327 pp. $8.00.

Certain sections are considerably revised or augmented.

Transactions of the Moscow Mathematical Society, vol. 12. English translation of *Trudy Moskovskogo Matematiceskogo Obscestva*, American Mathematical Society, Providence, R.I., 1965. iv+524 pp. $5.30.

Probability and its Engineering Uses, 2nd ed. By Thornton C. Fry. Van Nostrand, Princeton, N.J., 1965. 462 pp. $12.00 College edition, $15.00 Reference edition.

Radiative Transfer on Discrete Spaces. By Rudolph W. Preisendorfer. Pergamon Press, Long Island City, N.Y., 1965. xi+462 pp. $20.00.

Elemente der Funktionalanalysis. By L. A. Ljusternik and W. I. Sobolev, Akademie-Verlag, Berlin, 1965. xi+256 pp. DM 25.

The English translation of the Russian original was reviewed in the MONTHLY of November 1962.

Leçons d'analyse fonctionelle, 4th ed. By Frederic Riesz and Bela Sz.-Nagy, Akadémiai Kiadó, Budapest, 1965. viii+490 pp. $8.50.

Differential and Integral Calculus, 3rd ed. By Edmund Landau, translated by Melvin Hausner and Martin Davis. Chelsea, New York, 1965. 372 pp. $6.00.

Concepts from Tensor Analysis and Differential Geometry, 2nd ed. By Tracy Y. Thomas. Academic Press, New York, 1965. viii+178 pp. $7.50.

A chapter on Einstein-Riemann spaces has been added.

The Environment in Modern Physics, a Study in Relativistic Mechanics. By Clive W. Kilmister. American Elsevier, New York, 1965. viii+134 pp. $5.00.

A Survey of Modern Algebra, 3rd ed. By Garrett Birkhoff and Saunders MacLane. Macmillan, New York, 1965. x+437 pp. $8.50.

There are minor revisions throughout the book, major revisions in the material on Boolean algebra and lattices, and an introduction to tensor products has been added.

Mathematical Puzzles and Pastimes, 2nd ed. By Aaron Bakst. Van Nostrand, Princeton, N.J., 1965. 242 pp. $5.50.

Nijenrode Lectures—Asymptotic Distribution Modulo One. Edited by J. F. Koksma and L. Kuipers. Papers presented at the Nuffic International Summer Session in Science, sponsored by NATO, held in Breukelen, the Netherlands, in August, 1962. Noordhoff, Groningen, Holland, 1965. 203 pp. $2.90.

Calculus and Analytic Geometry, 2nd ed. By John F. Randolph. Wadsworth, Belmont, Calif., 1965. 640 pp. $10.95.

Theoretical and Mathematical Biology. Edited by Talbot H. Waterman and Harold J. Morowitz. Blaisdell, Waltham, Mass., 1965. 426 pp. $12.50.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor B. A. Amira, Hebrew University, Israel, Founder and Editor of the Journal d'Analyse Mathématique, was nominated Chevalier de la Légion d'Honneur by the Government of the French Republic in recognition of his activity in the field of International Cultural Relations.

Professor P. N. Carpenter, Grove City College, received an honorary Doctor of Science degree at Grove City College's Commencement on June 4, 1966.

Professor E. D. Ebert, University of Toledo, has been named a recipient of the University of Toledo's 1966 Outstanding Teacher Awards.

Professor Mark Kac, Rockefeller University, has been awarded an honorary degree of Doctor of Science by the Case Institute of Technology.

Mr. George Lenchner, Valley Stream North High School, Franklin Square, New York, has been awarded Harvard University's Distinguished Secondary School Teacher Award.

Professor Emeritus E. B. Mode, Boston University, represented the Association at the Investiture of Joseph Leo Driscoll as President of Southeastern Massachusetts Technological Institute on June 9, 1966.

Professor K. E. Whipple, Georgia State College, represented the Association at the inauguration of Harry M. Philpott as President of Auburn University on May 13, 1966.

University of Miami: Professor A. D. Wallace, University of Florida, has been appointed Professor; Dr. Marvin Mielke, Institute for Advanced Study, has been appointed Assistant Professor.

Wellesley College: Assistant Professor Torsten Norvig, University of Massachusetts, has been appointed Assistant Professor; Dr. Bernice L. Auslander has been promoted to Assistant Professor.

University of Wisconsin, Madison: Associate Professor A. A. Johnson, University of Toledo, and Assistant Professor J. G. Harvey, University of Illinois, have been appointed

Assistant Professors; Professors Steve Armentrout, State University of Iowa, and B. H. Neumann, Australian National University, have been appointed Visiting Professors; Associate Professors Anatole Beck and Fred Brauer have been promoted to Professors; Assistant Professors Lawrence Levy, P. E. Miles, and J. R. Smart have been promoted to Associate Professors; Associate Professor D. W. Crowe will be on leave for the fall semester at the Mathematical Institute of the Hungarian Academy of Sciences; Professor Walter Rudin will be on leave for the coming year at the University of California at La Jolla.

Associate Professor H. C. Kennedy, Providence College, is on sabbatical leave of absence for 1966–67 at the University of Turin, Italy, during the tenure of a Fulbright Research Grant.

Professor B. J. Lockhart, U. S. Naval Postgraduate School, has been appointed Dean of Curricula.

Assistant Professor Bernard Martin, Central Washington State College, has been promoted to Associate Professor.

Dr. H. D. Mills, International Business Machines, Rockville, Maryland, has been named Director of the IBM Federal Systems Division's Computer Mathematics Department.

Dr. G. F. Rose, System Development Corporation, Santa Monica, California, has been appointed Guest Professor at the Mathematisches Institut der Technischen Hochschule in Munich, Germany.

Assistant Professor Bobby Sanders, Texas Christian University, has been promoted to Associate Professor.

Assistant Professor J. M. Scandura, Florida State University, has been appointed Associate Professor at the Graduate School of Education of the University of Pennsylvania.

Professor L. F. Ollmann, Hofstra College, died on April 8, 1966. He was a member of the Association for 32 years.

Professor Emeritus Annie M. Pegram, Durham, North Carolina, died on April 13, 1966. She was a member of the Association for 47 years.

Professor Emeritus Anna Pell Wheeler, Bryn Mawr College, died on March 26, 1966. She was a charter member of the Association.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL MEETING OF THE ALLEGHENY MOUNTAIN SECTION

Seventy-five members and twenty-three guests were greeted by Waynesburg College President, Dr. B. M. Rich, at the fortieth annual meeting of the Section held in Waynesburg, Pa. April 30, 1966. Chairman F. E. Justis of Geneva College presided at the first morning session and conducted the business meeting. I. D. Peters of West Virginia University and Frank Perry of Washington and Jefferson College reported on the Annual Mathematics Contest. In West Virginia 3500 students from 77 schools participated and in Western Pennsylvania 5331 students from 116 schools. Six scored above 80.

C. V. Coffman of Carnegie Institute of Technology accepted gift memberships in MAA for his winning team in the William Lowell Putnam competition for the Section.

J. B. Bartoo, Penn State University, announced a memorial fund established at Penn State for the former governor, Evan Johnson, Jr. who died July 13, 1965. The term of governor is being completed by P. N. Carpenter, Grove City College. The chairman appointed A. B. Cunningham, West Virginia University, F. H. Steen, Allegheny College and George Laush, University of Pittsburgh to nominate a successor for the three year term beginning July 1, 1966.

The following were appointed by Professor Justis to study the CUPM curriculum proposals and to report at the 1967 meeting: C. F. Sebesta, Duquesne, Craig Comstock, Ward Bouwsma and Allan Krall, Penn State; Franz Hiergeist and Charles Church, West Virginia; Robert Berry, Washington and Jefferson; Anthony Pagano, Slippery Rock State, and George Novak, California State.

Dr. Bartoo reported on the committee formed to lobby in Harrisburg for stronger certification requirements for both elementary and secondary teachers. Other members of the committee are Richard Moore, Carnegie Institute of Technology, and Bernard Jacobson, Albert Filano and Martin Hubley of the Philadelphia Section.

Professor Peters presided at the session for short papers. Local arrangements were ably made by L. T. Moston, Waynesburg College, who presided at the afternoon session. An invitation from Professor Peters and West Virginia University to meet in Morgantown in connection with the University's 1967 centennial celebration was gratefully accepted. The date is May 6, 1967.

A group of portraits of eleven great mathematicians of the past made by Hwa S. Hahn, assistant professor of mathematics at Penn State University, was displayed in the coffee room.

A twenty-minute film on the 1964 Workshop at Stanford University, held by the Calculus Project of the Committee on Educational Media, was presented by Bruce Cornwall of the Calculus Project, Cleveland. O. F. H. Bert, charter member, reminisced on MAA's fifty years.

The following papers were given:

1. *Differential equations in the undergraduate curriculum*, by I. I. Kolodner, Carnegie Institute of Technology (by invitation).

2. *Computing and information*, by Preston Hammer, Penn State University (by invitation).

3. *The role of intuition*, by R. L. Wilder, president, MAA (by invitation).

4. *On asymptotic and convergent series*, by Craig Comstock, Penn State University.

The notion of asymptotic series is often treated as a converse of convergent series. In this note a single picture of both types of series is given, and conditions under which a series may be both convergent and asymptotic are shown.

5. *Properties of binary operators and their operator tables*, by George Reilly, Westinghouse Electric Corporation, Pittsburgh.

The operator tables for binary operators on a finite set of elements contain all information regarding the system. A computer algorithm has been developed for generation of Latin squares, loops, and tables for finite groups, which will identify which tables are group multiplicative systems. Some of the properties of the tables, considered as matrices, correspond to group properties.

6. *A generalization of the Mittag-Leffler function*, by A. M. Chak, West Virginia University.

The Mittag-Leffler function has been generalized by many authors in various directions. The purpose of this paper is to give another generalization and to study its properties by the help of operational calculus of two variables.

7. *Lattice-point solution of a problem in permutations*, by C. A. Church, Jr., and H. W. Gould, West Virginia University, presented by Professor Church.

The classical permutation problem of Terquem, following Riordan, can be stated as follows.

Consider combinations of n numbered things in natural (rising) order, with $f(n, r)$ the number of r-combinations with odd elements in odd position and even elements in even position, or, what is equivalent, with $f(n, r)$ the number of combinations with an equal number of odd and even elements for r even and with the number of odd elements one greater than the number of even for r odd.

In order to determine a recurrence for $f(n, r)$ and to find an explicit formula for $f(n, r)$ we consider a generalization modulo m due to Skolem (see Netto's *Kombinatorik*, 1927). We give a simple proof by enumeration of lattice paths for Skolem's formula for the more general problem.

BERTHA W. MATHER, *Secretary*

APRIL MEETING OF THE NEBRASKA SECTION

The forty-second annual meeting of the Nebraska Section of the MAA was held on April 30, 1966, at the Nebraska Center for Continuing Education, Lincoln, Nebraska, in conjunction with the seventy-sixth annual meeting of the Nebraska Academy of Sciences. Rev. E. A. Sharp, Chairman of the Section, presided. There were 65 persons present of whom 52 were members of the Association.

The following officers were elected for 1966–67: Chairman, Professor C. H. Frick, University of South Dakota; Vice-Chairman, Rev. E. A. Sharp, Creighton University; Secretary-Treasurer, Professor H. M. Cox, University of Nebraska. Professor J. M. Earl of The University of Omaha was continued as chairman of the Mathematics Contest Committee. Representatives of the Nebraska Section of the MAA, of the Nebraska Section of the National Council of Teachers of Mathematics, The Nebraska Actuaries Club, and the Nebraska Academy of Sciences are on the Contest Committee.

The following papers were presented:

1. *The Nebraska Mathematics Contest*, by J. M. Earl, University of Omaha, and H. M. Cox, University of Nebraska.

Some 4632 contest papers from 163 high schools were scored and reports printed by use of IBM 1230, IBM 534, and IBM 7040. The median team score was 59 as compared with 60 for the 1965 Contest. Reprints of the report printed in the *Nebraska Guidance Digest* for May 1966 are available upon request.

2. *A commutativity problem*, by E. P. Armendariz, University of Nebraska.

3. *Polynomials generated by a Bürmann series*, by O. G. Ruehr, University of Omaha.

Let $y(x)$ denote a formal power series possessing a formal inverse. The Bürmann series expansion of $y(kx)$ in powers of $y(x)$ generates a set of polynomials, $S_n(k)$. If $y(x)$ represents an entire function, sufficient conditions are given in the form of inequalities on the series coefficients of the inverse $x(y)$, so that the asymptotic behavior of $S_n(k)$ for large n is determined. The problem is related to that of determining the asymptotic behavior of the modified Faber polynomials associated with $x(y)$. Several examples were discussed and a coefficient problem was posed.

4. *Programmed learning in mathematics*, by H. L. Hunzeker, University of Omaha.

5. *Some results pertaining to sigma functions*, by J. T. Wallen, University of Nebraska.

Various problems concerning the sigma function were discussed and results given. Analogues were given for the unitary sigma function for which more complete results were obtained.

6. *Algebraic topology for undergraduates*, by A. B. Willcox, Amherst College and Vice President of the MAA (invited lecture).

7. *Turing machines, algorithms and their relation to computers*, by B. R. Gfeller, University of Omaha.

Description of Turing machines and what they do, relation of Turing machines to algorithms and the use of high speed computers.

8. *Difference methods of solving parabolic partial differential equations*, by Gangu Hingorani, Northern Natural Gas Company and Creighton University.

Parabolic differential equations can be solved as a system of simultaneous equations or by the grid method given initial and boundary conditions. Convergence and stability of difference schemes were discussed. Results were obtained for a particular heat conduction problem using a digital computer.

9. *Subbases applied to the existence of coefficient fields in commutative algebras*, by George Haddix, Creighton University.

Let F be a superfield of the field K. Then a subbasis of F over K is defined as a subset M of F such that K intersect M is empty where F equals $K(M)$ and where for any subset M' of M, m in $M-M'$ implies that the irreducible polynomial which m satisfies over K remains irreducible over $K(M')$. This concept can be applied to the problem of the existence of a coefficient field in a commutative algebra which contains a given field contained in the algebra.

10. *A characterization of the Jacobson radical*, by T. L. Jenkins, University of Nebraska.

11. *Direct sums of radical rings*, by A. E. Hoffman, University of Nebraska.

12. *Second order differential inequalities*, by K. W. Schrader, University of Nebraska.

13. *CUPM proposal for a general mathematics curriculum*, Panel discussion by H. L. Hunzeker, University of Omaha, D. Fuller, Creighton University, L. M. Larsen, Kearnsy State College, and W. A. Raab, University of South Dakota.

H. M. Cox, *Secretary*

APRIL MEETING OF THE OHIO SECTION

The fiftieth annual meeting of the Ohio Section of the MAA was held at Ohio Wesleyan University, Delaware, Ohio, on Saturday April 23, 1966. Professor William Fishback, Chairman of the Section, presided at the general and business sessions, and professors Dean Robb and Holbrook MacNeille presided at the program sessions. One hundred thirty registered including one hundred twenty-one members of the Association.

The following officers were elected: Chairman, Professor David Lipsich, University of Cincinnati; Chairman Elect, Professor Daniel Finkbeiner, Kenyon College; Secretary-Treasurer, Professor Foster Brooks, Kent State University.

Program Committee: Professor D. H. Staley, Ohio Wesleyan University; Professor B. J. Yozwiak, Youngstown University; Professor R. A. Clark, Case Institute of Technology.

Regular reports were made by the Chairman and the Secretary-Treasurer and special reports by Professor H. C. Trimble of The Ohio State University for the Committee on Teacher Training and Certification (CONTAC) and by Professor C. E. Capel of Miami University for the Committee on Curriculum (CONCUR).

At this fiftieth annual meeting of the Section special recognition was given at the luncheon to Professor V. B. Caris, retired, of The Ohio State University, only charter member of the Association present, and to Professor Rufus Crane, retired, of Ohio Wesleyan University, Secretary-Treasurer of the Ohio Section for twenty-one years (1926–1947), who was present with Mrs. Crane.

The following program was presented:

1. *The high costs of trying*, by W. T. Fishback, Ohio University, Athens, (Chairman's address).

The work of the Committees on Teacher Training and Accreditation and on Curriculum of the Section has ended its first phase successfully, but further action by them is needed if the goals of adequately prepared public school teachers and adequate preparation of incoming college students

are to be realized. Further Section activity seems in order in helping junior college teachers upgrade themselves and their programs, in maintaining adequate pregraduate programs in mathematics, and in controlling the growth of graduate programs in the state. This will require increased financial support for the Section and active participation of more Section members.

2. *Green's function for a spherical cell*, by H. W. Vayo, University of Toledo.

Electrical potential problems occurring in the electrophysiology of nerve and muscle differ from those usually considered in electrical theory for two reasons: first, the relation between current flow and polarization across a membrane of tissues is nonlinear, and second, the interior of the tissue is not strictly equipotential. We suppose that u_e and u_i are the steady state potentials out side and inside a tissue covered by a closed polarized membrane S. Then u_e and u_i are both harmonic functions which satisfy the nonlinear boundary condition at S. From these the Green's function is derived for a spherical cell.

3. *Report on a mathematics course for business students*, by L. C. Peck, Miami University.

Based on a minimum of $1\frac{1}{2}$ years of high school algebra, the year course covers linear programming through the simplex method, an introduction to matrices, a brief treatment of mathematics of investment, the derivative and integral for polynomial functions, and an introduction to probability.

Many results are intuitive generalizations. Applications are stressed.

Taught in large (200) sections with small problem sessions, major tests and the examination are multiple-choice, computer-graded. Part of output is a frequency distribution of scores, and an item analysis of the test.

4. *A theorem on linearly ordered maximal families*, by Nand Kishore, University of Toledo.

Let $X=\{x_1, x_2, \cdots, x_n\}$ be a set of n elements, and $P(X)$ be the power set of X which is partially ordered by set inclusion. Let S be a subset of $P(X)$ such that the nonempty elements of $P(X)-S$ are pairwise disjoint. And let r_k, $0\leqq r_k$, $1\leqq k<n$, denote the number of k-element sets of $P(X)-S$. Then there are

$$n! + \binom{r_1}{2}(n-2)! - \sum_{k=1}^{n-1} r_k k!\,(n-k)!$$

linearly ordered maximal families in S.

5. *Topology via geometry*, by J. L. Smith, Muskingum College.

Starting with congruences in Euclidean metric geometry, Klein's definition of geometry is used to view topology as a geometry by examining properties of a topology which remain invariant under the transformation of stretching.

6. *The digital design of dynamos*, by R. F. Jackson, University of Toledo.

The rational design of rotating electrical machines entails numerical solutions of Maxwell's equations in discontinuous structures of ferromagnetic materials. A lattice of points adequately defining the geometrical configuration requires storage of data exceeding the usual core memory of electronic computers unless one subdivides the cross-section into strips defined by the simplest possible lattices, joining the solutions by interpolation and making use of any periodic character of the cross section.

7. *A computer science program at The Ohio State University*, by T. W. Hildebrandt, The Ohio State University.

The Ohio State University has offered training in computer-related subjects and degree programs with computer emphasis in the Department of Mathematics for several years. This spring the Dean of the College of Arts and Sciences is establishing a Division of Computer Science which is expected to become a full-fledged department. This will permit the expansion and broadening of these programs in recognition of the fact that much of the present work with computers has taken on a nonnumeric character. This paper describes the considerations leading to the establish-

ment of the Division and indicates the proposed requirements and courses of study both for undergraduate and graduate degrees.

8. *Matrices over rings in which finitely generated ideals are principal*, by Melvin Henriksen, Case Institute of Technology.

Let $\mathfrak{A}_n$ denote the ring of $n \times n$ matrices over the commutative ring with identity $\mathfrak{A}$. We call $\mathfrak{A}$ an *F-ring* if each of its finitely generated ideals is principal, an *Hemite-Ring* if for every $A \in \mathfrak{A}_n$, there exists a nonsingular $Q \in \mathfrak{A}_n$ such that AQ is diagonal and we call $\mathfrak{A}$ an Elementary Division ring if for every such A, there exist nonsingular P, Q in $\mathfrak{A}_n$ such that PAQ is a diagonal matrix. The purpose of this talk is to gather together the known results about these rings and to pose open problems about them. See Kaplansky, TAMS, 66 (1949) 464–491; Gillman and Henriksen, TAMS, 82 (1956) 362–391, and Henriksen, Michigan Math. J., 3 (1956) 159–163.

9. *A group of calculus films recently produced by the Committee on Educational Media, MAA*, were shown by Holbrook MacNeille, Case Institute of Technology.

FOSTER BROOKS, *Secretary*

APRIL MEETING OF THE OKLAHOMA-ARKANSAS SECTION

The annual spring meeting of the Oklahoma-Arkansas Section was held at Oklahoma Baptist University, Shawnee, Oklahoma on April 1 and 2, 1966. There were 151 persons registered, of whom 107 were members of the Association.

At the business session Professor James Nickel reported that approximately 3800 high school students participated in the high school contest in Oklahoma and Arkansas; Professor Gene Levy, the Governor from our section reported on the activities of the Board of Governors; and Professor Glenn Haddock reported on the Sectional Officers meeting at Cornell last August.

Officers elected for 1966–67 were: Chairman James Scroggs of the University of Arkansas; Vice-Chairman Herbert Monks of Northeastern State College; Secretary-Treasurer Harold Huneke of The University of Oklahoma.

The spring meeting in 1967 will be held at Northeastern State College, Tahlequah, Oklahoma and the 1968 meeting will be held at the Federal Aviation Agency in Oklahoma City.

The Saturday morning session was devoted to the G.C.M.C. Report of the Committee on the Undergraduate Program. By invitation of the section, Professor G. Baley Price addressed the meeting on the CUPM recommendations on college mathematics offerings.

Following Dr. Price's report, a panel consisting of Professor Leslie Dwight of Southeastern State College, Professor Lysle Mason of Phillips University, and Professor R. B. Deal of Oklahoma State University discussed the topic: "Advantages and Disadvantages of the G.C.M.C. Program from my School's Viewpoint."

After discussion from the floor the following resolution was passed without dispute: "That the Section go on record recommending that any high school offering four years of mathematics should include in the last year the content of the course in Elementary Functions as recommended in the G.C.M.C. report." The Secretary was instructed to send notes of this resolution to teachers and administrators in Oklahoma and Arkansas.

The following papers were presented:

1. *An algorithm for computing polynomials of best approximation*, by J. E. Howland, The University of Oklahoma.

The problem considered in the paper is the approximation of a real-valued function by polynomials of degree $\leqq n$ such that the maximal deviation is as small as possible. For f defined on $[a,b]$, the author defines and characterizes polynomials of best approximation of degree $\leqq n$ to f on finite subsets of $[a, b]$, and develops a computational method of generating a sequence of poly-

nomials approximating f on finite subsets of $[a, b]$ which, if f is continuous on $[a, b]$, converges uniformly on $[a, b]$ to the polynomial of best approximation of degree $\leqq n$ to f on $[a, b]$.

2. *The generalized roulette in the plane*, by R. L. Simons, Oklahoma City University.

The general equation of a roulette, the locus of a point on a ray of a circle, as the circle is rolled along a general curve was derived. The special cases of the cycloid, epicycloid, etc., are obtained by suitable selection of the base curve.

3. *Paradoxes in locus problems*, by J. A. Nickel, Oklahoma City University.

(a) The locus of a point a given distance from a fixed point and (b) the locus of a point equally distant from two fixed points was considered by using the Minkowsky metric $D(A, B)=(|a_1-b_1|^p+|a_2-b_2|^p)^{1/p}$ as the distance function. The "circle" as first defined suggests a similarity of circular and rectangular neighborhoods. The "bisector" of the second locus is in general not a line. Letting p equal to one or tend to infinity produces exceptional cases which may take on any of three different forms.

4. *Fermat's Theorem—A geometrical approach*, by Ray Gazik, The University of Oklahoma.

It is shown that there are no nontrivial solutions in integers x, y, a, p for the equation $x^p+y^p=a^p$ if p is an odd prime and a is a power of a prime.

5. *A generalization of Filippov's Lemma*, by M. Q. Jacobs, The University of Oklahoma.

Let f be a continuous function, $F:R\times R^p\to R^q$. Let Ω be an upper semicontinuous mapping from R to the complete lattice of closed subsets of $R^p(=\mathcal{C}(R^p))$ such that $\Omega(t)$ is compact and nonempty for each $t\in R$. Define $F(t)=f(t, \Omega(t))$ for each $t\in R$. Let y be a μ-measurable function $y:R\to R^q$ such that $y(t)\in F(t)$ for each $t\in R$. Then there is a μ-measurable $u:R\to R^p$ such that $u(t)\in\Omega(t)$ for each $t\in R$ and $y(t)=f(t, u(t))$ for almost every $t\in R$.

This result reduces to Filippov's if the independent variable t is required to lie in a compact interval Γ and if $\bigcup_{t\in\Gamma}\Omega(t)$ is relatively compact.

6. *Characterization of an inverse arc map*, by John Jobe, Oklahoma State University.

If X is a space and f is a map such that $f(X)=Y$, then f is an inverse arc map if and only if for each arc L in Y there exists an arc L_1 in X such that $f(L_1)=L$. A factorization theorem analogous to one by G. T. Whyburn, *Analytic Topology*, Amer. Math. Soc., Colloq. Publ., 28 (1963), is stated in relation to an inverse arc map. A characterization of an inverse arc map is given.

7. *Mappings of T_1 topologies into topologies*, by Forrest Whitfield, Oklahoma State University.

Let S be a space and T a T_1 space with topologies σ and τ. Let g be a mapping of τ into σ. There exists a unique continuous mapping f of S into T such that $f^{-1}=g$ if and only if g satisfies the conditions (1), $g(\phi)=\phi$, (2), $g(\bigcup V_\alpha)=\bigcup g(V_\alpha)$, (3), $g\ (U\cap V)=g(U)\cap g(V)$, (4), $\bigcap g(V_\alpha)=0$, if $\bigcap V_\alpha=0$, and (5), if $N\in\sigma$, $N\subset g(V)$ for some V, where α belongs to an arbitrary indexing set.

8. *Some ergodic type problems*, by Professor Tetsundo Sikiguchi, University of Arkansas,

The following ergodic type problems are discussed. (a) Khinchin's theorem on continued fractions; (b) application of Hermann Weyl's theorem; (c) problem proposed by Dutch mathematician J. Duxemburg, "A necessary and sufficient condition that a set S of real numbers has Lebesgue measure zero is that all its subsets are Lebesgue measurable zero."

9. *The convex kernel of a class of noncontinuous functions*, by C. R. Smith, Oklahoma State University.

Given the class of noncontinuous functions mapping the unit interval into itself such that zero is a fixed point and each map has at most a finite number of discontinuities, we show that this set is not convex and that its convex kernel is nonempty. A characterization of this kernel is then given.

10. *A solution of the word problem for the braid group*, by J. H. Yates, Oklahoma State University.

The theory of braids was originated by E. Artin in 1925 to obtain structures similar in some

ways to knots only much simpler in concept. The main problem facing Artin was the classification or determination of equivalence of the braids. Artin solved this problem, by geometrical means in 1925 and again in 1947. Bohenenblust and Chow both succeeded in solving this geometrical problem by group theoretical means in 1947. In this paper a brief description of the braid groups is presented, along with a discussion of the solution of the word problem, which is the algebraic formulation of the classification problem of braids.

11. *An algebraic proof of Brouwer's Theorem on the degree of a map*, by John Keesee, The University of Arkansas.

Let K and L be triangulations of compact, orientable, n-dimensional manifolds M and N. Let $t:K\rightarrow L$ be a simplicial approximation of a map $f:M\rightarrow N$. Choose an n-simplex s^n of L and let m be the number of n-simplexes of K mapped by t onto s^n with orientation preserved less the number of n-simplexes mapped by t onto s^n with orientation reversed. Brouwer's fundamental theorem on the degree of a map states that the integer n depends only on the homotopy class of the map f. Most proofs of this result have a highly geometric flavor. The purpose here is to give a short, completely algebraic proof based on the topological invariance of the simplicial homology groups.

H. V. HUNEKE, *Secretary-Treasurer*

APRIL MEETING OF THE TEXAS SECTION

The annual spring meeting of the Texas Section of the MAA was held on the campus of Southern Methodist University, Dallas, Texas, on April 15–16, 1966. There were 287 persons registered including 210 members of the Association.

Papers were presented on both days in two concurrent sessions chaired by Professors J. T. Mohat of North Texas State University, A. D. Stewart of Prairie View A. and M. College, Bobby Sanders of Texas Christian University, Dale Maness of Austin College, and Father Theodosius Demen of Dallas University.

The welcoming address was given by Provost W. L. Ayres of Southern Methodist University at a dinner meeting on April 15, 1966 which was presided over by Professor C. J. Pipes of Southern Methodist University.

The general session on Saturday, April 16, 1966 was chaired by Professor C. R. Deeter of Texas Christian University and included reports of committees and of the secretary-treasurer. At this session the following officers were elected for the coming year: Chairman, C. J. Pipes of Southern Methodist University; Vice-Chairman, Dale Maness of Austin College; Secretary-Treasurer, Ben T. Goldbeck, Jr., of Texas Christian Unive sity.

The invited speakers for the meeting were Professor P. R. Halmos, University of Michigan and University of Miami, who spoke on "Dilation Theory" and Professor R. L. Wilder, University of Michigan, who spoke on "The Role of Intuition."

The following papers were presented:

1. *Discrete harmonic kernels*, by R. E. Huddleston, Texas Christian University.

The study of reproducing discrete harmonic kernels in a finite dimensional Hilbert space was introduced by Deeter and Springer in a paper in which many of the theorems on ordinary harmonic kernels were shown to have discrete analogues. Also they proved, using Fourier series methods, the convergence of the discrete harmonic kernel on a rectangular domain.

The purpose of this paper is to extend the domain of convergence of the discrete harmonic kernel to regions which are composed of finite unions of rectangles by using the general methods developed by Bramble and Hubbard.

2. *Extension of a ϕ-derivation*, by Tom Davis, Sam Houston State College.

Let R and R' be rings and ϕ a homomorphism of R into R'. Then D is a ϕ-derivation on R with values in R' iff (1). $D:R\rightarrow R'$; (2). $D(x+y)=D(x)+D(y)$, for every x, $y\in R$; (3). $D(xy)=D(x)y+\phi(x)D(y)$, for every x, $y\in R$. A proof is given that if $R\subseteq R'$, with R' a field and R a commutative integral domain, then D can be extended uniquely to the field of quotients of R.

3. *Bounded connectedness and Chebychev sets*, by Daniel Wulbert, University of Texas.

Let (X, d) be a metric space. A subset M of X is *boundedly connected* if M intersect S is connected for each open sphere S in X.

THEOREM: If M is a locally compact Chebychev set (each point in X admits a unique closest point in M) in a normed linear space, the following are equivalent: (A) M admits a continuous metric projection, (B) M is approximatively compact, (C) M is boundedly connected. THEOREM: A closed set M in a uniformly convex Banach space is convex if and only if in every equivalent norm topology M is a Chebychev set that admits a continuous metric projection. THEOREM: A compact Chebychev set in a normed linear space is the continuous image of $[0, 1]$.

4. *Near-rings with identities defined on cyclic groups*, by J. J. Malone, Jr., University of Houston, and J. R. Clay, George Washington University.

THEOREM: Let $(G, +)$ be a cyclic group. If $(G, +, {}^{\bullet})$ is a near-ring with identity, then $(G, +, {}^{\bullet})$ is a commutative ring with identity.

The theorem is proved by showing that the identity must be a generator of $(G, +)$.

5. *A proof of the decomposition theorem for a modular lattice with a least element*, by A. E. Borm, Southwest Texas State College.

Most books on lattice theory discuss the problem of direct product decompositions in terms of congruence relations. The following theorem is proved directly without involving the additional concept of a congruence relation: "A modular lattice L, with a least element 0, is isomorphic to the direct product of two nontrivial lattices if and only if it has two nontrivial ideals H_1 and H_2, such that $H_1 \wedge H_2 = \{0\}$, and $H_1 \vee H_2 = L$." The proof is straightforward. The main difficulty lies in showing that the representation $x = a \vee b$ of an element x of L as the join of an element a of H_1 and an element b of H_2, is unique.

6. *A characterization of reducibility for doubly stochastic matrices*, by R. D. Sinkhorn, University of Houston.

Denote the $n \times n$ identity matrix by I and the $n \times n$ matrix whose elements all equal $1/n$ by J. THEOREM: A necessary and sufficient condition that an $n \times n$ doubly stochastic matrix S be reducible is that $S-I-J$ be singular. If $\eta\ (S-I-J) = r-1$, then there exists a permutation P such that $P^T SP = S_1 \oplus S_2 \oplus \cdots \oplus S_r$ where each S_k, $k = 1, \cdots, r$, is doubly stochastic and irreducible.

7. *On binary sequences*, by G. S. Innis, Defense Research Laboratory, University of Texas.

Let $x_i = \pm 1$ for each i, $i = 1, 2, \cdots n$, and let $C_k = \sum_1^{n-k} x_i x_{i+k}$. Let $p = \max_{1 \le k \le n-1} |C_k|$. Such a sequence is called a code of length n and type p. It is known that codes of type 1, Barker codes, do not exceed length 13. In some applications a longer code of minimum type is desired and in others the ratio $n:p$ is of primary interest. In this paper some preliminary results for codes of type p for $p > 1$ are presented.

8. *On a singular Abel-type integral equation on E^2*, by A. D. Stewart, Prairie View A. and M. College.

9. *Matrix summability of sequences of bounded variation*, by D. F. Dawson, North Texas State University.

THEOREM 1. If a matrix A sums every sequence of bounded variation, in the sense of Brannen and Kuttner, then A sums a convergent sequence not of bounded variation. THEOREM 1 is a special case of (but was instrumental in proving) the following result. THEOREM 2. If M is a countable collection of matrices, each of which sums every sequence of bounded variation, then there exists a convergent sequence not of a bounded variation which every matrix in M sums.

10. *Some structure theorems on compact commutative integral domains with characteristic p*, by J. E. Cude, University of Texas.

Kaplansky has shown that a compact integral domain with an identity has an open Jacobson

radical and that the ring modulo the radical is a division ring. THEOREM: If A is a compact commutative integral domain with an identity and having characteristic p, then there exists a maximal subfield F of A such that F is isomorphic to $A/J(A)$, where $J(A)$ is the Jacobson radical of A. Furthermore, F is the unique maximal subfield of A.

11. *A theorem on polynomial identities*, by T. P. Kezlan, University of Texas.

Let f be any polynomial in (noncommuting) indeterminates $x_1, \cdots, x_n$ with integral coefficients (at least one of which is the integer 1). THEOREM: The following are equivalent: (1)Every semi-simple ring satisfying $f\equiv 0$ is commutative. (ii) Every ring satisfying $f\equiv 0$ has nil commutator ideal. (iii) For every prime p the ring of 2×2 matrices over the prime field with p elements does *not* satisfy $f\equiv 0$.

12. *A first order iteration process for simultaneous equations*, by H. A. Luther and W. F. Stewart, Texas A. and M. University.

The judgment of an iterative solution of simultaneous equations properly entails more than the rate of convergence. An item of concern is whether or not we can arrive at a solution point whose approximate location is known.

Let the "region of influence" of a given solution point be the set of all starting values which (for a given iterative technique) have that solution point as the limiting value of the iteration.

The present procedure was devised as part of a program for studying methods of altering the regions of influence rather than rates of convergence.

The equations considered can be linear or nonlinear. By using first order partial derivatives, there is created a technique which has the Newton-Raphson quality of converging to every solution point whose Jacobian is nonvanishing. The method is computationally simpler than the Newton-Raphson method; it is, however, first order rather than second order. An arbitrary function can be used parametrically.

13. *Convolution and Fourier series*, by Jack Bryant, Texas A. and M. University.

Consider the convolution $g*h$ of functions g and h in $L(0, 2\pi)$. The main result of this paper is a theorem extending a theorem of Salem on writing a given function as the convolution of other functions. Our result gives very general conditions on a Banach space B of functions on $[0, 2\pi]$ which imply that every f in B can be written $f=g*h$, where g is in L and h is in B. We apply our methods to the problem of convergence almost everywhere of Fourier series.

14. *An embedding theorem for topological semigroups*, by G. E. Mattingly, Sam Houston State College.

Necessary and sufficient conditions are given for embedding a topological semigroup as a neighborhood of one of its points in a topological group. A corollary extends a previously known result.

15. *On the set of idempotents in a convex cone*, by R. V. McPherson, University of Texas.

A convex cone is a set with two operations, addition and scalar multiplication, which satisfy the axioms for a vector space except that the scalars are restricted to the set of nonnegative real numbers and the cancellation law for addition is not necessarily satisfied. This allows cones to have nonzero idempotents. There is a natural generalization of linear variety to convex cones. The set of nonzero idempotents in a convex cone is either void or a variety.

16. *Interlocking sequences on completely regular spaces*, by Peter Morris, University of Texas.

A sequence $\{A_n, B_n\}$ of pairs of disjoint, nonvoid, closed subsets of a topological space X is an *Interlocking Sequence* (*I.S.*) if every set of the form $\bigcap \{C_n: C_n=A_n \text{ or } C_n=B_n, n=1, 2, \cdots, N\} \neq \phi$. Let $C^*(X)$ denote the Banach space (sup norm) of real, bounded, continuous, functions on X. THEOREM: Let X be a completely regular Hausdorff space. Consider the following: (i) X admits an *I.S.*; (ii) $C^*(X)$ contains l_1; (iii) $C^*(X)$ contains $C([0, 1])$; (iv) X can be mapped continuously onto a dense subset of $[0, 1]$; (v) βX is not dispersed. Then (ii) through (v) are equivalent and imply (i). If X is normal, (i) through (v) are equivalent.

17. *Periodic solutions of linear differential equations*, by Louis Brand, University of Houston.

The linear differential equation of order n, $P(D)x=f(t)$, where $P(D)$ is a polynomial in $D, f(t)$ a piecewise continuous function of period p, always has a solution of period p and class C^{n-1} if the zeros of $P(s)$ are not poles of $F(s)=\mathfrak{L}\{f(t)\}$. If its Laplace transform is $P(s)x=Q(s)+F(s)$, the initial values $x(0)=\alpha$, $x'(0)=\beta$, $\cdots$ that produce the wave are determined by the n equations $Q^{(i)}(s_j)+F^{(i)}(s_j)=0$, $i=0, 1, \cdots, m_j-1$, where s_j is a zero of $P(s)$ of multiplicity m_j.

18. *Quotients of the space* c_0, by H. E. Lacey, University of Texas.

The space c_0 is the Banach space of all sequences of real numbers converging to 0. The purpose of this note is to prove that a quotient space of c_0 is either finite dimensional or has the property that each indefinite dimensional closed subspace contains a copy of c_0.

19. *The identification of fundamental truth functions by decimal arithmetization*, by T. J. Bradshaw, Texas Christian University.

20. *Formulas of propositional logic containing countably many variables*, by A. Ullman, Texas Christian University.

A mapping is defined on logical formulas of countably many variables into subsets of the Cantor ternary set. Topological characterizations of various classes of logical formulas can be given in terms of the associated subsets of the Cantor ternary set.

21. *On a substructure in certain semirings*, by R. G. Dean, Arlington State College.

Analogous to an ideal, a multiplicative substructure, a cell, is defined for semirings. Certain semirings in which addition distributes over addition are discussed. In a mutually distributive semiring S, cosets modulo a cell C, form a semiring S/C, a homomorphic image of S. Also, the inverse image of an all element under a homomorphism ϕ, is a cell C, and the homomorphic image of S, $S\phi$, is unity-semi-isomorphic to S/C. A mapping which reverses operations is discussed. An example is given to distinguish difference semirings modulo an ideal from quotient semirings modulo a cell.

22. *The lattice of congruences on an inverse semigroup*, by H. E. Scheiblich, University of Texas.

THEOREM: Let S be an inverse semigroup with E as its set of idempotents and λ its lattice of congruences. Define $\theta=\{(\rho_1, \rho_2)\rho_1, \rho_2\in\lambda; e\rho_1\cap E=e\rho_2\cap E \text{ for each } e\in E\}$. Then (i) θ is a congruence on λ, (ii) each θ class is a complete modular lattice, and (iii) the natural homomorphism of λ onto λ/θ is a complete lattice homomorphism.

23. *Nonnegative matrices with doubly stochastic powers*, by Nancy Bedford, University of Houston.

It is shown that a nonnegative irreducible matrix A which is not doubly stochastic has a doubly stochastic power A^m if and only if (1) A is cyclic of index h with $(m, h)>1$, and (2) there exist $s_i>0$, $i=1, \cdots, h$, such that the $n_i \times n_{i+1}$ matrices A_i in the Frobenius normal form are s_i/s_{i+1} stochastic, all indices modulo h; the matrices A_i^T are $(n_i/n_{i+1})(s_i/s_{i+1})$ stochastic; $s_i=s_{i+m}$; and there is at least one index j with $s_j/s_{j+1}\neq 1$ and / or $(n_j/n_{j+1})(s_j/s_{j+1})\neq 1$. These results are applied to reducible matrices.

24. *The abstract nature of computer programs*, by John Reynolds, Texas Christian University.

The abstract nature of computer programs seems to indicate that they form closed algebraic systems. A study of the intrinsic elements of computer programs could lead to more efficient methods of hardware design. This study deserves investigation for the purposes of both applied and pure mathematics.

25. *Number systems with irrational bases*, by R. R. Bunten, 6570th USAF Personnel Research Laboratory.

Most numeric operations are performed using numbers to the base 10, but systems based on six, twelve, and twenty are still in use. Modern electronic computers use binary and octal numbers. By using the same idea and some of the same operations, any number can serve as the

base of a number system. This paper examines the implications of using an irrational number as a base, with particular attention to the number e, which allows the expression,

$$e^3 + 2e^2 + e + 1 + e^{-1} + 2e^{-2} + e^{-3},$$

to be written 1211.121_e.

26. *Isomorphisms and homeomorphisms*, by A. Ullman, Texas Christian University.

A general definition of isomorphism is given and it is shown that the usual algebraic isomorphisms are particular cases of this definition. Similarly, it is shown that the definition of a topological homeomorphism is a particular case.

27. *Matrices, conics, and quadrics*, by A. R. Amir-Moez, Texas Technological College.

The equation $ax^2+2bxy+cy^2=0$ can be written as $XAX'=0$. Here $X=(xy1)$. If we write the equation of the conic in homogeneous form, we get a quadratic form in x, y, and z, and then A is the symmetric matrix of this quadratic form. Suppose that Q is the matrix of the quadratic form $ax^2+2bxy+cy^2$. Then points of intersection of the conic and the line $X=X_0+tD$ will be obtained from the equation $(DQD')t^2+2(X_0AD')t+X_0AX_0'=0$, where $X_0=(x_0\ y_0\ 1)$ and $D=(r\ s\ 0)$. Through this second degree equation one can discuss many ideas related to a conic, such as center, axes, vertices, $\cdots$, without change of coordinate system. These ideas can easily be generalized for quadrics in higher-dimensional spaces. Details can be found in "Matrix techniques, trigonometry, and analytic geometry," Edwards Brothers Inc., Ann Arbor, Michigan.

B. T. GOLDBECK, JR., *Secretary-Treasurer*

MAY MEETING OF THE ILLINOIS SECTION

The forty-fifth annual meeting of the Illinois Section of the MAA was held at Saint Dominic College, St. Charles, Illinois, on May 13 and 14, 1966. Amos Black, Section Chairman, presided at the afternoon session and the dinner meeting Friday evening, and the incoming Section Chairman, Wayne McGaughey, presided at the Saturday morning session. Approximately 115 persons, including 88 members of the Association, attended one or more of the sessions and/or the dinner meeting.

The following officers were elected: Chairman, A. Wayne McGaughey, Bradley University; Vice-Chairman, R. D. Boswell, Monmouth College; Secretary-Treasurer, Arnold Wendt, Western Illinois University.

The business meeting which followed the first session on Friday afternoon opened with general announcements by the secretary and the treasurer's report, which showed a decrease of $268.67 in the treasury balance for the year. The primary reason for the decrease was a donation of $200 to the National MAA Contest Committee. The Illinois Section MAA contest continues to be the major source of funds which support other Section activities.

Professor G. S. Young, Vice-President of MAA, then reported on items of interest to the members. In the future are new programs for improving collegiate instruction in science and mathematics. CUPM is now also concerned about qualified staff, especially at the junior college level. Partly due to CUPM efforts, the number of collegiate institutions requiring no mathematics of prospective elementary teachers continues to decrease, and six semester hours of mathematics is now a common requirement for elementary teachers.

James Beach, chairman of the Membership Committee, reported that each junior college Mathematics Department chairman was contacted and encouraged to attend the annual meeting. He also announced that the Departments of Mathematics and Education at Northern Illinois University will sponsor a follow-up meeting to the 1962 CUPM meeting on the pre-service training of elementary teachers. The meeting will be held on Oct. 1, 1966, at the University Center at Northern Illinois University. Each Mathematics Department and Education Department is encouraged to send representatives.

Douglas Bey, chairman of the Secondary School Lecturer Committee reported that for the five years preceding the 1965–66 academic year the Committee received a grant for secondary mathematics lecturers directly from NSF, and 50–100 schools per year were served. In 1965–66, however, the NSF made the grant for all secondary science and mathematics lecturers to the State Academy of Science. Because of these altered circumstances, only twenty-eight requests for mathematics lecturers were received and only eighteen of these were assigned lecturers because of lack of funds. The State Academy has requested funds for two-hundred lecturers next year, and the mathematics share should be larger than this year.

Walter McCurdy, chairman of the Contest Committee, reported another successful year. Three-hundred-forty-two schools and 20,757 students participated. In view of the ample profit shown by the contest, a motion to grant $200 to the National MAA Contest Committee, if donations by the Committee are solicited, was duly made, seconded, and approved.

Arnold Wendt, chairman of the Committee on the Strengthening of the Teaching of Mathematics, reported two main activities during the year: CSTM submitted recommendations to the Governor's Task Force on Education and to the Teacher Education and Professional Standards Commission of the Illinois Education Association. The recommendations dealt primarily with upgrading certification procedures and standards, especially at the elementary, junior high, and junior college levels.

Hiram Paley, chairman of the Undergraduate Participation Committee reported on the Committee's project of encouraging participation in the Putnam competition by offering cash prizes and/or MAA memberships to the top fifteen Illinois contestants. The Section appropriated $200 to continue this activity next year.

The Section accepted invitations from the University of Illinois and Southern Illinois University at Edwardsville to hold the Section meeting at these institutions in 1967 and 1968, respectively.

Professor Franz Hohn gave the dinner address. He gave an illustrated lecture on his experiences as a visiting lecturer in Holland.

After a cordial welcome from Sister Mary Paul, O.P., Ph.D., President of Saint Dominic College, the following program was presented during the Friday afternoon and Saturday morning sessions:

1. *On the periodicity of the sum of two periodic functions*, by C. G. Townsend, Southern Illinois University.

THEOREM. *If f and g are periodic functions, $f(d)$ a positive absolute maximum and d a point of continuity of f, then $f+g$ is periodic only if p/q is rational for some period p of f and some period q of g.*

The sum of two periodic functions f and g is always periodic if p/q is rational for some period p of f and some period q of g. Examples are available which show that the converse is not true in general. The above theorem specifies the weakest conditions the author has been able to find which insure that the converse does hold.

2. *Some hypergeometric identities*, by Gus DiAntonia, Northern Illinois University.

Three hypergeometric identities involving $F(a, b, c, z)$, are proved. By using the hypergeometric differential equation as a reduction formula on a, b, c, several other identities are deduced. These are closely related to the Gaussian contiguous relations. Further identities may be obtained by employing the Kummer solutions in their regions of intersections of the z plane. All of these carry over directly to the generalized hypergeometric differential equation.

3. *Some problems in cluster set theory*, by G. S. Young, Tulane University (by invitation).

4. *Partially ordered topological spaces, trees, and local trees*, by G. E. Dimitroff, Knox College. A sequence of theorems is presented which leads to several characterizations of trees and

local trees as partially ordered topological spaces having certain natural restrictions. The characterizations are viewed as descriptions of the inherent order properties possessed by trees and by local trees.

5. *A class of linear sequence spaces*, by Sister M. Catharina Bereiter, Siena Heights College.

The linear space $L(S)$ is defined by requiring that all subsequences with index sequence in S form absolutely convergent series. If S covers the natural numbers, the supremum of these sums forms a norm in the space $B(S)$ of sequences for which it is finite, and $B(S)$ is a Banach Space. Under easily satisfied conditions $L(S)$ and $B(S)$ differ from spaces such as (l_1) and (m). For certain S, "defined by a counting function Ω," $B(S)$ equals $L(S)$ and is separable, but not reflexive. The dual space can be given by sequences determined directly from Ω.

6. *A mnemonic simplification in linear algebra*, by Zamir Bavel, Southern Illinois University.

Let U and V be finite dimensional vector spaces with D and D' ordered bases for U, R and R' ordered bases for V, $\alpha \in U$, and $T: U \to V$ a linear transformation. Denote by $[\alpha]_D$ the *coordinate matrix of α relative to D*; by ${}^R[T]^D$ the *matrix representing T relative to D and R*; and by $P(D \to D')$ the *transition matrix from D to D'*. Also regard $D \to D'$ as instructions for substitution: To perform the *forward* substitution $D \to D'$, replace D' by D, and to perform the *backward* substitution $D \to D'$, replace D by D'. In either case, the substitution "consumes" the transition matrix. RULE: (a) A transition matrix adjacent to ${}^R[T]^D$ appears on the domain (alt. range) side of ${}^R[T]^D$ when the change-of-basis occurs in the domain (alt. range) space. (b) Perform a *forward* substitution in what *follows* a transition matrix; perform a *backward* substitution in what is *behind* a transition matrix. It is now easy to remember, prove, and "invent" such theorems as

$$P(D \to C)^C[T]^A P(A \to E) P(E \to B)[\alpha]_B = {}^D[T]^B[\alpha]_B = [T(\alpha)]_D,$$

since it is impossible to misstate them.

7. *A Problem in Elementary Set Theory*, (Hour Address), by Philip Dwinger, University of Illinois, Chicago Circle.

ARNOLD WENDT, *Secretary-Treasurer*

MAY MEETING OF THE INDIANA SECTION

The Indiana Section of the MAA met on Saturday, May 14, 1966, at Indiana State University, Terre Haute, in joint session with the Indiana Council of Teachers of Mathematics. Approximately 200 persons attended, of whom 70 were members of the Association. Chairman George Springer of Indiana University presided, and President A. C. Rankin of Indiana State University welcomed the participants.

At its business meeting, the Section elected the following officers for the year 1966–67: Robert Zink, Purdue University, Chairman; Kenneth Sidebottom, Indiana Central College, Vice-Chairman; George Pedrick, Purdue University, Secretary-Treasurer. The Section also voted to give special recognition annually to the top Indiana team and individual in the Putnam Competition and directed its Executive Committee to determine a tangible expression of this recognition. A communication from the Mathematics Department of Indiana University was read in which it was announced that the department "each year reserves one of its regular stipends for graduate study for the participant in the Putnam Competition who ranks highest among the contestants in the State of Indiana."

The program consisted of three hour lectures as follows:

1. *Rotations, angles and trigonometry*, by R. J. Troyer, Dartmouth College.

2. *Puzzles, platonism and extraversion*, by E. E. Moise, Harvard University.

3. *Continued fractions in stability theory*, by J. S. Frame, Michigan State University.

PAUL MIELKE, *Secretary*

MAY MEETING OF THE WISCONSIN SECTION

The annual meeting of the Wisconsin Section of the MAA was held at Wisconsin State University, Eau Claire, on May 7, 1966. Professor Harold Glander, Chairman of the Section, presided. There were 201 persons in attendance including 72 members of the Association.

Mr. Owen White, Manitowoc County Center of the University of Wisconsin, presented a report of the Contest Committee. From 340 high schools, 23,273 registered for the preliminary contest held in local high schools on Thursday, February 10; 1123 students from 226 schools participated in the final contest which was held in 27 centers located throughout the state of Wisconsin on Saturday, March 26.

Professor John V. Finch, Beloit College, Governor of the MAA from the Wisconsin Section, gave a report. The following officers were elected to serve during the school year of 1966–67: Chairman: Sister Madeleine Sophie of Alverno College; Vice-Chairman: Professor E. F. Wilde, Beloit College; Secretary-Treasurer: L. F. Wahlstrom, Wisconsin State University, Eau Claire.

The following papers were presented:

1. *Rolle's Theorem in abstract space*, by Morris Marden, University of Wisconsin-Milwaukee.

As background for the extension of Rolle's Theorem in abstract space, the speaker first surveyed the corresponding results in the complex plane. These results included the theorems of Gauss and Lucas on the location of the zeros of the derivative of a polynomial and then the theorems of Laguerre and Brocher which are invariant under linear transformation. Generalizations of the latter two theorems to a vector field over an algebraically closed field were then stated as recently developed by Hörmander and by Marden.

2. *Differential rings and a Bruns Theorem for the harmonic oscillator*, by Lawrence Markus, University of Minnesota.

A differential ring is an algebraic ring with a derivation, that is, a mapping into itself satisfying the usual laws for derivatives of sums and products. The constants, whose derivatives are zero, form a subring. The main results of formal integration theory can be stated in this context. For example, the standard derivation on the ring of real exponential-trigonometric polynomials is surjective. Another application concerns the recognition of the ring of all first integrals of a dynamical system, as in the Bruns Theorem of celestial mechanics. An analogous result is proved for the harmonic oscillator.

3. *Math movies—what for?*, by Seymour Schuster, University of Minnesota.

The speaker presented a discussion of the weaknesses of existing mathematical films as suitable teaching instruments and the probable reasons for these weaknesses. He discussed the philosophy of the College Geometry Project (of Minnemath) toward integrating films with independent-study materials. Also discussed were the attitudes of the College Geometry Project staff toward overcoming the above-mentioned weaknesses; namely, to produce films with greater involvement on the part of mathematicians and artists. Two geometry films were viewed and discussed as attempts to realize and fulfill the aims of the College Geometry Project.

4. *Foundations of probability and statistics*, by Arthur Copeland, Sr., University of Michigan.

This theory of the foundations results from the conviction that the probability of the occurrence of an event should be related to the occurrence and that the relation should be evident from the definition of probability. It is shown that the usual definitions fail to indicate this relation. A new definition clarifying the relation is based on the testing of probabilities. Since the usual theory of statistical tests conceals the relation between probabilities and occurrences, a revised theory is presented. The revised theory gives a clearer understanding of why experimental science is so successful.

L. F. WAHLSTROM, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

Forty-eighth Summer Meeting, University of Toronto, Toronto, Ontario, Canada, August 28–30, 1967.

ALLEGHENY MOUNTAIN, West Virginia University, Morgantown, West Virginia, May 6, 1967.

ILLINOIS, University of Illinois, Urbana, May 14–15, 1967.

INDIANA, Purdue University, Lafayette, November 5, 1966.

IOWA, Drake University, Des Moines, April 21, 1967.

KANSAS, Fort Hays State College, Hays, April 22, 1967.

KENTUCKY, Murray State University, Murray, Spring 1967.

LOUISIANA-MISSISSIPPI, Jung Hotel, New Orleans, Louisiana, March 4–5, 1967.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN

MINNESOTA

MISSOURI, Northeast Missouri State Teachers College, Kirksville, April 29, 1967.

NEBRASKA, University of South Dakota, Vermillion, May 6, 1967.

NEW JERSEY, Rutgers, The State University, New Brunswick, November 12, 1966.

NORTHEASTERN, Trinity College, Hartford, Connecticut, November 26, 1966.

NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.

OHIO

OKLAHOMA-ARKANSAS, Northeastern State College, Tahlequah, Oklahoma, March–April, 1967.

PACIFIC NORTHWEST, University of Montana, Missoula, June 16–17, 1967.

PHILADELPHIA, Villanova University, Villanova, November 19, 1966.

ROCKY MOUNTAIN

SOUTHEASTERN, Florida Presbyterian College, St. Petersburg, Florida. March 31–April 1, 1967.

SOUTHERN CALIFORNIA, San Diego State College, San Diego, March 11, 1967.

SOUTHWESTERN, University of Arizona, Tucson, March 31–April 1, 1967.

TEXAS, Austin College, Sherman, April 14–15, 1967.

UPPER NEW YORK STATE, State University College, Plattsburgh, May 20, 1967.

WISCONSIN, St. Norbert College, DePere, May 6, 1967.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.

AMERICAN MATHEMATICAL SOCIETY, Houston, Texas, January 24–27, 1967.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Michigan State University, June 19–23, 1967.

ASSOCIATION FOR COMPUTING MACHINERY, Sheraton-Park, Washington, D. C., August 29–31, 1967.

ASSOCIATION FOR SYMBOLIC LOGIC, Houston, Texas, January 23–24, 1967.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Indianapolis, November 24–26, 1966.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Houston, Texas, January 28, 1967.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

THE AMERICAN MATHEMATICAL MONTHLY, Vol. 73, No. 8, October, 1966, pp. 817-936

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 9

CONTENTS

NOVEMBER 1966

THE AMERICAN MATHEMATICAL MONTHLY

(Founded in 1894 by Benjamin F. Finkel)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. FICKEN, Department of Mathematics, New York University, New York, N. Y. 10453. Manuscripts of papers should be submitted to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE TO AUTHORS

The MONTHLY welcomes papers presenting valid mathematics, of rather general interest, at a level intelligible to persons with two years of full-time graduate study. Some novelty of content, viewpoint, or arrangement is essential. Expository articles are particularly desired. State the context and the principal aim of the paper early. Address yourself quite explicitly to the reader described above, communicating your ideas to him clearly and attractively.

The title should be brief and meaningful. Since the title will be quoted and reproduced by laymen, it should contain no symbols unfamiliar to laymen.

Articles should be typewritten, double-spaced, on $8\frac{1}{2}\times 11''$ paper of very good quality. Submit the original (and a duplicate if convenient) keeping a complete copy for yourself. To avoid loss and delay notify us of any change of address.

The typescript should be prepared with extreme care. Misprints are highly obnoxious; so are dangling participles. Put name and address between title and text. Put references at the end with bracketed citations in the text. Avoid footnotes; instead, use clearly designated remarks in the text. Put acknowledgments at the very end, just before the bibliography. Be generous with spacing and displays. *Keep notation simple.* For a matrix the notation $[a_{jk}]$ is recommended, with det $[a_{jk}]$ or $|a_{jk}|$ for the determinant. On doubtful questions regarding format or notation, observe practices in current issues of the MONTHLY, or consult the applicable sections of the "Author's Manual" of the American Mathematical Society.

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June-July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

INTRINSIC METRICS

HERMAN GLUCK, Harvard University and The University of Pennsylvania

1. Introduction. Let M be a set and d: $M \times M \to R$ a real-valued function satisfying

(1) $d(x, y) \geqq 0$; $d(x, y) = 0$ iff $x = y$ (positive definiteness)

(2) $d(x, y) = d(y, x)$ (symmetry)

(3) $d(x, z) \leqq d(x, y) + d(y, z)$ (triangle inequality).

Such a function d is said to be a *metric* on M, and the pair (M, d) is then called a *metric space*. Certain metrics on a set M seem to be "more natural" than others. For example, let M be the set of complex numbers of unit length, and define a metric d on M by

$$d(z, z') = |z - z'|.$$

The set M is simply the unit circle in the complex plane and the metric d is "inherited" by M from the ordinary metric in the plane. Now consider the points -1 and 1 in M. Their distance from one another is 2, yet every path in M from -1 to 1 has length at least π. The metric d would hardly be a practical metric for people living and moving about in M. The situation could be improved, however, by increasing the metric d to a metric d_0 for which $d_0(z, z')$ is the length (in the metric d) of the shorter arc from z to z'. This new metric is topologically equivalent to the old one, and assigns the same length to curves in M as did d. Furthermore, it is now possible to travel from z to z' in M by a path of length $d_0(z, z')$, and it is in this sense that the metric d_0 is "better" or "more natural" than the metric d. The metric d_0 is an example of an *intrinsic metric* on M (the precise definition will be given later).

Spaces with intrinsic metrics have been studied extensively in the literature [**1, 2**]. My object here is to examine in detail some of the ideas leading up to the notion of an intrinsic metric. Because of the elementary character of this paper, a number of proofs will be left to the reader.

2. Finitely compact metric spaces. A metric space (M, d) has the *Bolzano-Weierstrass property* if every bounded infinite set in M has a limit point in M. It has the *Heine-Borel property* if every closed and bounded subset of M is compact. These two properties are in fact equivalent, and a metric space having them is said to be *finitely compact.* A finitely compact metric space is separable, locally compact and complete. Conversely, any locally compact separable metric space possesses an equivalent metric in terms of which it is finitely compact [**3**].

3. Nearness. We are going to need a notion of "nearness" for points in a metric space (M, d) which is independent of the particular metric d and depends only on the topology which d induces on M. A typical way of developing such a notion is via *admissible coverings* of M. A covering is *admissible* if every point of M lies in the interior of some member of the covering. Then, relative to such a

covering, two points in M will be "near" if there is some member of the covering containing them both. Notice, however, that what counts is not the particular covering but rather the "nearness relation" which it induces on $M\times M$. So we are led to a definition of the following sort. (This approach will remind the reader of that used in the study of uniform structures.)

DEFINITION 3.1. *Let M be a topological space. A sleeve is a subset Σ of $M\times M$ which is symmetric (i.e., $(x, y)\in\Sigma$ iff $(y, x)\in\Sigma$) and is also a neighborhood of the diagonal $\{(x, x): x\in M\}$, of $M\times M$.*

The points x, y in M will be "near" relative to the sleeve Σ if $(x, y)\in\Sigma$. A subset of a sleeve which is itself a sleeve will be called a *subsleeve*. Notice that the intersection of two sleeves is again a sleeve. Hence any two sleeves in $M\times M$ always have a common subsleeve. A sleeve Σ which is a closed subset of $M\times M$ will be called a *closed sleeve.*

Given the sleeve Σ, we define a sleeve Σ^2 as follows: $(x, y)\in\Sigma^2$ iff there is a point $u\in M$ such that $(x, u)\in\Sigma$ and $(u, y)\in\Sigma$. In a like manner we can define the sleeves Σ^3, Σ^4, $\cdots$, and we will have

$$\Sigma = \Sigma^1 \subset \Sigma^2 \subset \Sigma^3 \subset \Sigma^4 \subset \cdots$$

Carrying this to an extreme, define

$$\Sigma^\infty = \bigcup_{n=1}^{\infty} \Sigma^n.$$

Then $(x, y)\in\Sigma^\infty$ iff there is a finite sequence of points in M,

$$x = m_0, m_1, m_2, \cdots, m_r = y,$$

such that for each $i=1, 2, \cdots, r$, $(m_{i-1}, m_i)\in\Sigma$. Such a finite sequence of points will be called a *Σ-chain* from x to y in M. A sleeve Σ will be said to be *wide* if $\Sigma^\infty=M\times M$, that is, if every pair of points in M can be joined by a Σ-chain. The proof of the following theorem is immediate.

THEOREM 3.2. *Every sleeve in $M\times M$ is wide iff M is connected.*

Suppose now that (M, d) is a metric space. Then for each number $\delta>0$, consider the closed sleeve $\Sigma_\delta=\{(x, y): d(x, y)\leqq\delta\}$. If Σ is a sleeve in $M\times M$ and C is a compact subset of M, there is always a Lebesgue number $\delta=\delta(C, \Sigma) >0$, such that $\Sigma_\delta\cap(C\times C)\subset\Sigma$. That is, if $x, y\in C$ and $d(x, y)\leqq\delta$, then $(x, y)\in\Sigma$.

4. Construction of the metric d_Σ. Let (M, d) be a metric space and Σ a sleeve in $M\times M$. If $x=m_0, m_1, m_2, \cdots, m_r=y$ is a Σ-chain in M from x to y, then the number

$$\sum_{i=1}^{r} d(m_{i-1}, m_i)$$

will be referred to as the *length* of this Σ-chain. For the remainder of this section we will assume that Σ is a wide sleeve, so that every pair of points in M can be joined by at least one Σ-chain. This will automatically be the case, according to Theorem 3.2, if M is connected.

DEFINITION 4.1. *Let $d_\Sigma(x, y)$ denote the greatest lower bound of the set of lengths of all Σ-chains in M from x to y.*

Picturesquely speaking, $d_\Sigma(x, y)$ is the practical distance from x to y for those people in M whose strides are "subordinate" to the sleeve Σ.

REMARKS. (1) $d_\Sigma(x, y) \geqq d(x, y)$ for any $x, y \in M$. We write: $d_\Sigma \geqq d$.

(2) If $(x, y) \in \Sigma$, then $d_\Sigma(x, y) = d(x, y)$. Hence the functions d_Σ and d agree for points "sufficiently" close to one another.

(3) If C is a compact subset of M, let $\delta = \delta(C, \Sigma)$ be a Lebesgue number for C relative to Σ. Then for $x, y \in C$ with $d(x, y) \leqq \delta$, we have $d_\Sigma(x, y) = d(x, y)$. So for $x, y \in C$, $d_\Sigma(x, y) \leqq \delta$ iff $d(x, y) \leqq \delta$, in which case the two are equal.

(4) If $\Sigma = \Sigma_\delta$, then instead of writing d_{Σ_δ}, we will write simply d_δ.

THEOREM 4.2. *d_Σ: $M \times M \to R$ is a metric on M equivalent to d.*

Positive definiteness follows from Remark (1) above. Symmetry and the triangle inequality follow directly from the definition of d_Σ. So d_Σ is a metric on M. It is equivalent to d by Remark (2) above.

THEOREM 4.3. *If (M, d) is finitely compact, then so is (M, d_Σ).*

Let C be closed and bounded in (M, d_Σ). Then C is closed in (M, d) by the preceding theorem, and bounded in (M, d) by Remark (1) above. Hence C is compact in (M, d). Again by the preceding theorem, C must also be compact in (M, d_Σ), so (M, d_Σ) is finitely compact.

REMARK. The converse to this theorem is false. Let M be the set of real numbers and d the metric on M defined by

$$d(x, y) = \frac{|x - y|}{1 + |x - y|}.$$

The metric d is equivalent to the ordinary absolute value metric but is bounded, so (M, d) is certainly not finitely compact. On the other hand, if $\Sigma = \Sigma_{1/2}$, it is easily checked that $(M, d_{1/2})$ is finitely compact.

THEOREM 4.4. *If (M, d) is complete, then so is (M, d_Σ).*

Let (x_n) be a Cauchy sequence in (M, d_Σ). Since $d_\Sigma \geqq d$, the sequence (x_n) is also Cauchy in (M, d), where it must therefore converge. Since the metrics d and d_Σ are equivalent, (x_n) also converges in (M, d_Σ), and the theorem follows.

REMARK. The converse to this theorem is also false. Let

$$M = \left\{(x, y) : y = \sin\frac{1}{x}, x > 0\right\} \subset R^2,$$

and let d be the metric on M induced by the inclusion $M \subset R^2$. Then (M, d) is not complete, because it is not closed in R^2. But if we define the sleeve $\Sigma \subset M \times M$ by

$$\Sigma = \left\{\left(\left(x, \sin\frac{1}{x}\right), \left(x', \sin\frac{1}{x'}\right)\right) : \left|\frac{1}{x} - \frac{1}{x'}\right| < 1\right\},$$

then d_Σ "stretches" M enough to make (M, d_Σ) complete.

According to the definition of $d_\Sigma(x, y)$, there is for each $\epsilon > 0$ a Σ-chain in M from x to y whose length is less than $d_\Sigma(x, y) + \epsilon$, but in many cases there will be no Σ-chain from x to y whose length is exactly $d_\Sigma(x, y)$. We seek now conditions on the metric space (M, d) and on the sleeve Σ which will guarantee the existence of a Σ-chain from x to y with length exactly $d_\Sigma(x, y)$.

Let (M, d) be a finitely compact metric space and $\Sigma \subset M \times M$ a wide sleeve. Fix the points $x, y \in M$ and consider the set

$$C = \{(x' \in M : d(x, x') \leqq 2d_\Sigma(x, y)\}.$$

Since (M, d) is finitely compact, C must be compact. Let $\delta = \delta(C, \Sigma) > 0$ be a Lebesgue number for C relative to Σ. Let N be an integer such that $N\delta > d_\Sigma(x, y)$.

LEMMA 4.5. *Under the above conditions, $d_\Sigma(x, y)$ is the greatest lower bound of the set of lengths of all Σ-chains in M from x to y which have at most $2N$ terms.*

$d_\Sigma(x, y)$ is the greatest lower bound of the set of lengths of all Σ-chains in M from x to y with length $\leqq \min (N\delta, 2d_\Sigma(x, y))$. Hence it will be sufficient to produce, for each such Σ-chain, a corresponding Σ-chain in M from x to y with at most $2N$ terms, whose length is no larger.

Let

$$x = m_0, m_1, m_2, \cdots, m_r = y$$

be a Σ-chain in M from x to y with length $\leqq \min (N\delta, 2d_\Sigma(x, y))$. First observe that, since $d(x, m_i) \leqq 2d_\Sigma(x, y)$, all the m_i lie in the compact set C. If this Σ-chain has more than $2N$ terms, then $r \geqq 2N$ and we can write $[d(m_0, m_1) + d(m_1, m_2)] + \cdots + [d(m_{2N-2}, m_{2N-1}) + d(m_{2N-1}, m_{2N})] \leqq N\delta$, in which case there must be an odd index i such that

$$d(m_{i-1}, m_i) + d(m_i, m_{i+1}) \leqq \delta.$$

But then $d(m_{i-1}, m_{i+1}) \leqq \delta$, so that $(m_{i-1}, m_{i+1}) \in \Sigma$. Therefore m_i can legitimately be dropped from the original Σ-chain to get a new Σ-chain from x to y with one term less and total length no larger. Repetition of this procedure then completes the proof of the lemma.

THEOREM 4.6. *If* (M, d) *is finitely compact and* Σ *is a closed sleeve, then for each* $x, y \in M$ *there will exist a* Σ*-chain from* x *to* y *with length exactly* $d_\Sigma(x, y)$.

Fix $x, y \in M$ and let C, δ and N be defined as above. For each $n = 1, 2, 3, \cdots$, let

$$C_n: x = m_{0n}, m_{1n}, m_{2n}, \cdots, m_{2N-2,n}, m_{2N-1,n} = y$$

be a Σ-chain from x to y with exactly $2N$ terms and with length $\leqq \min(d_\Sigma(x, y) + 1/n, 2d_\Sigma(x, y))$. The existence of the C_n is guaranteed by the preceding lemma. As before, all the m_{in} lie in the compact set C. Hence by repeatedly taking subsequences of (C_n), we can finally get a subsequence (C_{n_j}) such that (m_{i,n_j}) converges, say to m_i, for each $i = 0, 1, \cdots, 2N-1$. The limit sequence, $C_\infty: x = m_0, m_1, \cdots, m_{2N-1} = y$, will be a Σ-chain because Σ is a closed sleeve, and its length will be exactly $d_\Sigma(x, y)$, completing the proof.

REMARK. Theorem 4.6 is no longer true when finite compactness is replaced by completeness. Let $M = \{a, b, c_1, c_2, c_3, \cdots\}$, and let

$$d(a, b) = 1, \quad d(c_i, c_j) = 1 - \delta_{ij}, \quad d(a, c_i) = d(c_i, b) = 1/2 + 1/2^i.$$

For $x \neq y$, $\frac{1}{2} < d(x, y) \leqq 1$, so the triangle inequality is automatically satisfied. All Cauchy sequences are eventually constant, so the space (M, d) is certainly complete. On the other hand, if $\frac{1}{2} < \delta < 1$ then $d_\delta(a, b) = 1$, but although the sleeve Σ_δ is closed, there is no δ-chain from a to b with length exactly 1.

THEOREM 4.7. *Arc length is the same in the metrics* d *and* d_Σ.

C is a path from x to y in (M, d) iff it is a path from x to y in (M, d_Σ), because d and d_Σ are equivalent metrics. Furthermore, since C is a compact set, there is a $\delta > 0$ such that d and d_Σ agree for points of C closer than δ. It then follows immediately that C is rectifiable in (M, d) iff it is rectifiable in (M, d_Σ), and that d and d_Σ then assign the same arc length to C.

THEOREM 4.8. *Let* d *and* d^* *be metrics on the same set* M, *with* $d^* \geqq d$, *and let* Σ *be a wide sleeve in* $(M, d) \times (M, d)$. *Then* Σ *is also a wide sleeve in* $(M, d^*) \times (M, d^*)$, *and* $(d^*)_\Sigma \geqq d_\Sigma$.

Since $d^* \geqq d$, the identity map from $(M, d^*) \times (M, d^*)$ to $(M, d) \times (M, d)$ must be continuous. Hence Σ is a neighborhood of the diagonal in $(M, d^*) \times (M, d^*)$. Since the remaining conditions are clearly satisfied, Σ is a wide sleeve in $(M, d^*) \times (M, d^*)$. Any Σ-chain in (M, d^*) is also a Σ-chain in (M, d), and its length in (M, d^*) is greater than or equal to its length in (M, d). From this it immediately follows that $(d^*)_\Sigma \geqq d_\Sigma$.

THEOREM 4.9. *Let* (M, d) *be connected and let* Σ *and* Σ' *be sleeves in* $M \times M$.
(1) *If* $\Sigma' \subset \Sigma$, *then* $d_{\Sigma'} \geqq d_\Sigma$.
(2) *If* $\Sigma' \subset \Sigma$, *then* $(d_\Sigma)_{\Sigma'} = (d_{\Sigma'})_\Sigma = d_{\Sigma'}$.
(3) *In general*, $(d_\Sigma)_{\Sigma'} \neq (d_{\Sigma'})_\Sigma$.

(4) *In the following list, each metric is greater than or equal to all those listed on lower levels, while those on the same level can not in general be compared.*

$$\begin{array}{ccc} & d_{\Sigma\cap\Sigma'} & \\ (d_\Sigma)_{\Sigma'} & & (d_{\Sigma'})_\Sigma \\ d_\Sigma & & d_{\Sigma'} \\ & d & \end{array}$$

Proof left to the reader.

5. Construction of the metric d_0. From now on we will assume that the metric space (M, d) is connected. Hence each sleeve $\Sigma\subset M\times M$ is wide, and therefore the metric d_Σ is always well-defined.

For each pair of points x, $y\in M$, consider the set of numbers $\{d_\Sigma(x, y)\}_\Sigma$, where Σ ranges over all possible sleeves in $M\times M$. This set of numbers may or may not be bounded from above.

DEFINITION 5.1. *If, for each x, $y\in M$, the set $\{d_\Sigma(x, y)\}_\Sigma$ is bounded from above, then we say that the metric d is rectifiable, and define d_0: $M\times M\rightarrow R$ by $d_0(x, y)=\mathrm{LUB}_\Sigma d_\Sigma(x, y)$. We also say that d_0 is obtained from d by the process of rectification.*

Intuitively speaking, $d_0(x, y)$ is the smallest number which can serve as a practical distance from x to y for all people in M, regardless of their strides.

REMARKS. (1) If d is rectifiable, then $d_0\geqq d_\Sigma\geqq d$ for all sleeves Σ.

(2) If between each x, $y\in M$ there is a rectifiable path with length $L(x, y)$, then d is rectifiable and $d_0(x, y)\leqq L(x, y)$.

(3) In computing $\mathrm{LUB}_\Sigma d_\Sigma(x, y)$, one can restrict Σ to be a subsleeve of a given sleeve without altering the answer.

(4) One can also restrict Σ to be a closed sleeve without altering the answer, since $M\times M$ is a normal space.

It is easy to give examples of metrics which are not rectifiable. If $M=\{(x, y): y=x\sin 1/x \text{ for } x>0, y=0 \text{ for } x=0\}$ and d is the metric induced on M by the inclusion $M\subset R^2$, then d is not rectifiable in the sense of the above definition.

THEOREM 5.2. *If (M, d) is finitely compact and d is rectifiable, then*

$$d_0(x, y) = \mathrm{LUB}_\Sigma\, d_\Sigma(x, y) = \mathrm{LUB}_{\delta>0}\, d_\delta(x, y) = \lim_{\delta\to 0} d_\delta(x, y).$$

This is generally false, however, if (M, d) is not finitely compact.

Proof left to the reader.

THEOREM 5.3. *If d is rectifiable, then d_0 is a metric on M.*

Positive definiteness follows from the fact that $d_0\geqq d$. Symmetry follows

from the symmetry of each d_Σ. As for the triangle inequality, consider $d_0(x, z)$ and note that for each $\epsilon>0$ there is a sleeve Σ such that $d_0(x, z)<d_\Sigma(x, z)+\epsilon$. But $d_\Sigma(x, z)\leqq d_\Sigma(x, y)+d_\Sigma(y, z)\leqq d_0(x, y)+d_0(x, z)$. Hence $d_0(x, z)<d_0(x, y)+d_0(y, z)+\epsilon$. Since this holds for each $\epsilon>0$, the triangle inequality follows immediately.

THEOREM 5.4. *If d is rectifiable, then* (1) *d: $(M, d_0)\times(M, d_0)\to R$ is uniformly continuous, and* (2) *the identity map i: $(M, d_0)\to(M, d)$ is uniformly continuous.*

Proof trivial, because $d_0\geqq d$.

Even if d is rectifiable, d and d_0 need not be equivalent metrics. We construct a compact subset M of the plane, as follows. (See Figure 1.)

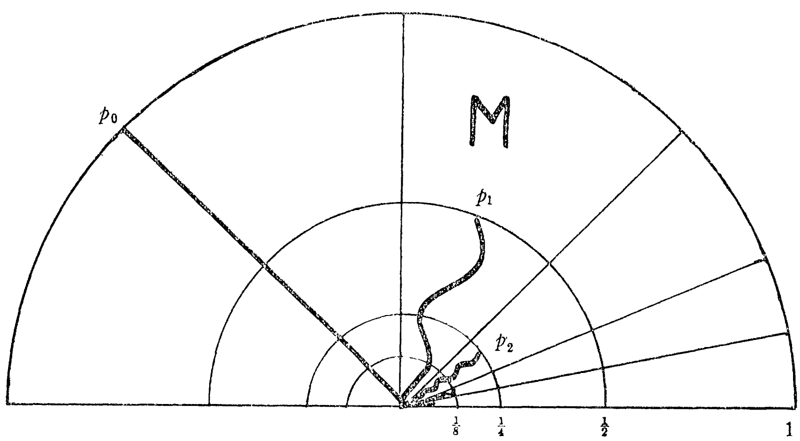


FIG. 1

In the sector: $0\leqq r\leqq 1$, $\pi/2<\phi<\pi$, draw an arc of length 1 from the origin to the point p_0: $r=1$, $\phi=3\pi/4$. In the sector: $0\leqq r\leqq\frac{1}{2}$, $\pi/4<\phi<\pi/2$, draw an arc of length 1 from the origin to the point p_1: $r=\frac{1}{2}$, $\phi=3\pi/8$. And in general, in the sector: $0\leqq r\leqq\frac{1}{2}^n$, $\pi/2^{n+1}<\phi<\pi/2^n$, draw an arc of length 1 from the origin to the point p_n: $r=\frac{1}{2}^n$, $\phi=3\pi/2^{n+2}$. The union of this sequence of arcs is the compact set M.

The metric d on M is induced by the inclusion $M\subset R^2$. Since every pair of points in M can be joined by a path of finite length, the metric d is rectifiable. Since $d(p_n, 0)=\frac{1}{2}^n$, the sequence $p_0, p_1, p_2, \cdots$ converges to 0 in (M, d). But $d_0(p_n, 0)=1$ for all n, so this sequence does not converge to 0 in (M, d_0). Hence the metrics d and d_0 are not equivalent.

REMARKS. *Relative to the above example*, observe that

(1) (M, d) is compact, yet (M, d_0) is not compact, nor even finitely compact.

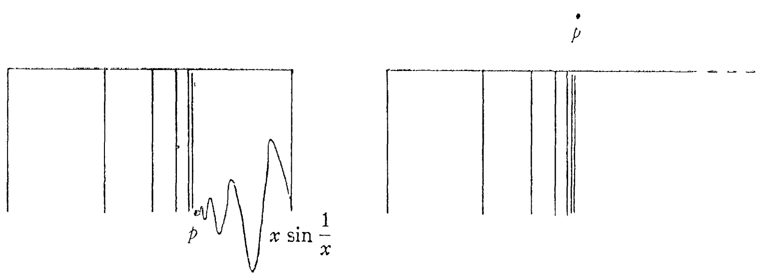
(2) (M, d) is complete, and so is (M, d_0).

(3) (M, d) is arcwise connected, and so is (M, d_0).

THEOREM 5.5. *If* (M, d) *is complete and* d *is rectifiable, then* (M, d_0) *is also complete. The converse is false.*

Proof left to the reader.

Consider the subset M of the plane given in Figure 2a below. If d is the metric on M induced by the inclusion $M \subset R^2$, then (M, d) is arcwise connected and d is rectifiable. (M, d_0), however, is not even connected! While (M, d_0) is not isometric to any subset of the plane, it is at least *homeomorphic* to the set indicated in Figure 2b.



2a 2b

FIG. 2

I want to prove now, in analogy to Theorem 4.7, that arc length is the same in the metrics d and d_0. A little caution must be exercised, however. If C is a curve in (M, d_0), then C is also a curve in (M, d) because the identity map $i\colon (M, d_0) \to (M, d)$ is continuous. The converse is false. Look in fact at the example following Theorem 5.4, in which the metrics d and d_0 are not equivalent. Let C be the compact curve which starts at p_0 and runs to 0 along the first arc, then out to p_1 and back to 0 along the second arc, then out to p_2 and back to 0 along the third arc, and so forth, ending finally back at 0. C is a curve in (M, d) but is not a curve in (M, d_0)! The statement of the following theorem is therefore slightly more complicated than that of its prototype, Theorem 4.7.

THEOREM 5.6. *Let* d *be a rectifiable metric on* M. *If the curve* $C \subset (M, d)$ *has length* L *in* (M, d), *then* C *is a curve in* (M, d_0) *and also has length* L *there. Conversely, if the curve* $C \subset (M, d_0)$ *has length* L *in* (M, d_0), *then* C *is a curve in* (M, d) *and has length* L *there.*

Without loss of generality, we may assume that C is a compact curve in (M, d) and let

$$f\colon [0, L] \to (M, d)$$

be a parametrization of C by arc length. Then, by Remark (2) in Section 5, we have

$$d_0(f(t), f(t')) \leqq |t - t'|,$$

from which it follows that f must also be continuous with respect to the metric d_0. Thus C is a curve in (M, d_0). The above inequality also implies that the length of C in (M, d_0) is at most L. But it is also at least L, because $d_0 \geqq d$, so its length in (M, d_0) is exactly L.

A curve in (M, d_0) is automatically a curve in (M, d), and the converse then follows immediately.

By analogy with Theorem 4.8, we state

THEOREM 5.7. *Let d and d^* be metrics on the same set M, with $d^* \geqq d$. Let (M, d^*), and therefore also (M, d), be connected. If the metric d^* is rectifiable, so is d, and then $(d^*)_0 \geqq d_0$.*

Each sleeve $\Sigma \subset (M, d) \times (M, d)$ is also a sleeve in $(M, d^*) \times (M, d^*)$, and then $(d^*)_\Sigma \geqq d_\Sigma$, according to Theorem 4.8. Hence $(d^*)_0 \geqq d_\Sigma$ for each such Σ. But then, by the very definition of rectification, d must be rectifiable and satisfy $(d^*)_0 \geqq d_0$.

THEOREM 5.8. *If d is a rectifiable metric on M, then so is each d_Σ, and $(d_\Sigma)_0 = d_0$.*

First of all, d and d_Σ are equivalent metrics on M, according to Theorem 4.2. Therefore sleeves in $(M, d) \times (M, d)$ coincide with sleeves in $(M, d_\Sigma) \times (M, d_\Sigma)$. Now pick $x, y \in M$ and consider $\text{LUB}_{\Sigma'}(d_\Sigma)_{\Sigma'}(x, y)$, if this exists. By Remark (3) in Section 5, we can restrict Σ' to be a subsleeve of Σ without altering the value of this LUB. But if $\Sigma' \subset \Sigma$, we know from Theorem 4.9 that $(d_\Sigma)_{\Sigma'} = d_{\Sigma'}$. Hence

$$(d_\Sigma)_0(x, y) = \text{LUB}_{\Sigma'}(d_\Sigma)_{\Sigma'}(x, y) = \text{LUB}_{\Sigma'}\, d_{\Sigma'}(x, y) = d_0(x, y),$$

and the theorem is proved.

THEOREM 5.9. *If (M, d) is finitely compact and d is rectifiable, then for each $x, y \in M$ there is a rectifiable arc from x to y in (M, d) with length $d_0(x, y)$.*

According to Theorem 5.2, $d_0(x, y) = \lim_{\delta \to 0} d_\delta(x, y)$. Also, since (M, d) is finitely compact and Σ_δ is a closed sleeve, Theorem 4.6 asserts that for each $\delta > 0$ there is a Σ_δ-chain from x to y with length exactly $d_\delta(x, y)$.

Now fix $x, y \in M$. For each $\delta > 0$, let

$$C_\delta: x = m_{0\delta}, m_{1\delta}, m_{2\delta}, \cdots, m_{r_\delta \delta} = y$$

be a Σ_δ-chain from x to y with length exactly $d_\delta(x, y)$. Let

$$L = d_0(x, y) \geqq d_\delta(x, y),$$

and note that each $m_{i\delta}$ lies in the compact set $C = \{x' \in M: d(x, x') \leqq L\}$.

Construct a map f_δ: $[0, L] \to (M, d)$ as follows. Let

$$\begin{aligned}
t_{0\delta} &= 0 \\
t_{1\delta} &= d(m_{0\delta}, m_{1\delta}) \\
t_{2\delta} &= d(m_{0\delta}, m_{1\delta}) + d(m_{1\delta}, m_{2\delta}) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
t_{r_\delta \delta} &= d(m_{0\delta}, m_{1\delta}) + \cdots + d(m_{r_\delta - 1, \delta}, m_{r_\delta \delta}) = d_\delta(x, y) \leqq L.
\end{aligned}$$

Then define f_δ by

$$\begin{aligned}
&f_\delta\colon [t_{i\delta}, t_{i+1,\delta}) \to m_{i\delta}, \quad \text{for } i = 0, 1, 2, \cdots, r_\delta - 1 \\
&f_\delta\colon [t_{r_\delta \delta}, L] \to m_{r_\delta \delta} = y.
\end{aligned}$$

The function f_δ is a step function with simple jump discontinuities at times $t_{1\delta}, t_{2\delta}, \cdots, t_{r_\delta \delta}$ (presuming that all the $m_{i\delta}$ are distinct), and satisfies the growth law

$$d(f_\delta(t), f_\delta(t')) \leqq |t - t'| + \delta.$$

Roughly speaking, the function f_δ is just another way of describing the chain C_δ.

For each $\delta > 0$, the image of f_δ lies in the fixed compact set C. Hence there is a sequence (f_{δ_j}) of these functions, with δ_j decreasing to 0, which converges pointwise to a function f on a countable dense subset D of $[0, L]$. For $t, t' \in D$, we have

$$d(f(t), f(t')) \leqq |t - t'| + \delta$$

for all $\delta > 0$, and hence

$$d(f(t), f(t')) \leqq |t - t'|.$$

Hence f is uniformly continuous and therefore has a unique extension f over $[0, L]$, satisfying the above growth law for all $t, t' \in [0, L]$. The uniformly continuous function f: $[0, L] \to (M, d)$ defines a path of length at most L connecting x with y. On the other hand, its length is at least L by Remark 2 on page 10, so its length is exactly L. This path is in fact an arc (i.e., without self-intersections), because otherwise removal of a loop would produce a path from x to y with length less than L, again contrary to Remark (2) in Section 5.

This completes the proof of the theorem. (The idea for this proof is due to my former student at Harvard, Lawrence Brown.)

COROLLARY. *If (M, d) is finitely compact and d is rectifiable, then both (M, d) and (M, d_0) must be arcwise connected.*

This is because a rectifiable arc in (M, d) is also a rectifiable arc in (M, d_0).

THEOREM 5.10. *Let (M, d) be finitely compact and d rectifiable. Then d_0 is also rectifiable, and $(d_0)_0 = d_0$.*

According to Theorem 5.9, each pair of points $x, y \in M$ can be connected by

a rectifiable arc in (M, d) of length $d_0(x, y)$. According to Theorem 5.6, this rectifiable arc is also a rectifiable arc in (M, d_0) with the same length, $d_0(x, y)$. But then Remark (2) in Section 5 implies that d_0 is rectifiable and that $(d_0)_0(x, y) \leqq d_0(x, y)$. Since $(d_0)_0 \geqq d_0$, the theorem follows.

REMARK. The example given on page 944 shows that the above theorem is no longer true upon deletion of the finite compactness condition. In that example, (M, d_0) was not connected, so that d_0 could not be rectified.

6. Faithful rectification.

DEFINITION 6.1. *If the metric d is rectifiable and d_0 is equivalent to d, then we say that d is faithfully rectifiable, and refer to the process as faithful rectification.*

THEOREM 6.2. *The following conditions are equivalent*: (1) *d is faithfully rectifiable.* (2) *d is rectifiable and d_0: $(M, d)\times(M, d)\to R$ is continuous.* (3) *d is rectifiable and the identity map i: $(M, d)\to(M, d_0)$ is continuous.*

Proof left to the reader.

THEOREM 6.3. *Let d be a faithfully rectifiable metric on M. For each pair of points $x, y\in M$ and each $\epsilon>0$, there is a sleeve Σ and a $\delta>0$ such that*

$$\left| d_0(x, y) - d_{\Sigma'}(x', y') \right| < \epsilon$$

for all sleeves $\Sigma'\subset\Sigma$ and points $x', y'\in M$ satisfying $d(x, x')<\delta$, $d(y, y')<\delta$.

Choose Σ so that $|d_0(x, y)-d_\Sigma(x, y)|<\epsilon/3$. Using the fact that d is faithfully rectifiable, choose $\delta>0$ so that $d_0(x, x')<\epsilon/3$ and $d_0(y, y')<\epsilon/3$ whenever $d(x, x')<\delta$ and $d(y, y')<\delta$. Then

$$\begin{aligned} \left| d_0(x, y) - d_{\Sigma'}(x', y') \right| &\leqq \left| d_0(x, y) - d_{\Sigma'}(x, y) \right| + \left| d_{\Sigma'}(x, y) - d_{\Sigma'}(x', y') \right| \\ &\leqq \left| d_0(x, y) - d_\Sigma(x, y) \right| + d_{\Sigma'}(x, x') + d_{\Sigma'}(y, y') \\ &\leqq \left| d_0(x, y) - d_\Sigma(x, y) \right| + d_0(x, x') + d_0(y, y') \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon, \end{aligned}$$

whenever $\Sigma'\subset\Sigma$ and $d(x, x')<\delta$, $d(y, y')<\delta$. Note that we have made use of part (1) of Theorem 4.9 to conclude that $|d_0(x, y)-d_{\Sigma'}(x, y)| \leqq |d_0(x, y) -d_\Sigma(x, y)|$.

THEOREM 6.4. *If d is rectifiable and (M, d_0) is compact, then d is automatically faithfully rectifiable.*

For then the identity map i: $(M, d_0)\to(M, d)$, which in any case is a continuous bijection, must be a homeomorphism. But then its inverse must be continuous, and so the theorem follows from part (3) of Theorem 6.2.

THEOREM 6.5. *If d is faithfully rectifiable and (M, d) is finitely compact, then (M, d_0) is also finitely compact.*

Let C be closed and bounded in (M, d_0). Then C is closed in (M, d) because d and d_0 are equivalent metrics, and bounded in (M, d) because $d \leqq d_0$. Hence C is compact in (M, d). But then again by the equivalence of d and d_0, C must be compact in (M, d_0), and the theorem follows.

REMARK. If d is faithfully rectifiable and (M, d_0) is finitely compact, it does not follow that (M, d) is finitely compact. The example mentioned in the Remark following Theorem 4.3 works equally well here, for if

$$d(x, y) = \frac{|x - y|}{1 + |x - y|},$$

then $d_0(x, y) = |x - y|$, so d_0 is equivalent to d. Then (M, d_0) is finitely compact, while (M, d) isn't.

The following question asks whether Theorem 5.10 remains true upon replacement of the hypotheses: "(M, d) is finitely compact and d is rectifiable" by the hypothesis: "d is faithfully rectifiable."

QUESTION. *Let d be a faithfully rectifiable metric on M. Is d_0 rectifiable, and does $(d_0)_0 = d_0$?*

7. Intrinsic metrics.

DEFINITION 7.1. *The metric d on M will be said to be intrinsic if d is rectifiable and $d_0 = d$ (so in fact, d must be faithfully rectifiable). Equivalently, d is intrinsic if (M, d) is connected and $d_\Sigma = d$ for all sleeves $\Sigma \subset M \times M$.*

THEOREM 7.2. *If (M, d) is finitely compact and d is rectifiable, then d_0 is an intrinsic metric on M.*

This is just a restatement of Theorem 5.10.

THEOREM 7.3. *Let (M, d) be finitely compact and d an intrinsic metric. Then for each $x, y \in M$ there is a rectifiable arc from x to y in (M, d) with length $d(x, y)$.*

This follows immediately from Theorem 5.9.

In a finitely compact space with an intrinsic metric, the distance between two points actually represents the length of the shortest path between them. The space is therefore quite reasonable to live in!

Consider the following

PROBLEM FOR METRIC SPACES. *Given a metric space (M, d), find an intrinsic metric $d^* > d$, which is equivalent to d.*

This problem is connected with the rectification process as follows.

THEOREM 7.4. *Let (M, d) be a metric space, and suppose there exists on M an intrinsic metric $d^* \geqq d$.*

(1) *Then d is rectifiable and $d_0 \leqq d^*$.*

(2) *If d^* is equivalent to d, then so must d_0 be equivalent to d. That is, d is faithfully rectifiable.*

(3) *Independent of the assumption in part* (2), *if* (M, d) *is finitely compact, then d_0 is intrinsic.*

Since d^* is an intrinsic metric, (M, d^*) is certainly connected. Then by Theorem 5.7, d must be rectifiable and satisfy $d_0 \leqq (d^*)_0 = d^*$, proving part (1).

If d^* is equivalent to d, consider the following diagram of identity maps:

$$\begin{tikzcd} & (M, d^*) \arrow[dl, "j"'] \arrow[dd, "k"] \\ (M, d_0) \arrow[dr, "i"'] & \\ & (M, d) \end{tikzcd}$$

The maps i, j and k are continuous. Since d^* is equivalent to d, k is a homeomorphism. But then $i^{-1} = jk^{-1}$ is continuous and, by Theorem 6.2, d_0 is equivalent to d, proving part (2).

Since d is rectifiable, part (3) follows from Theorem 7.2.

COROLLARY. *The above Problem for Metric Spaces has a solution for the finitely compact space* (M, d) *if and only if d is faithfully rectifiable, in which case d_0 is the minimum solution.*

REMARK. There are plenty of examples of spaces (M, d) for which the problem has no solution, since a necessary condition for the existence of a solution is that d be faithfully rectifiable. If the answer to the question posed at the end of Section 6 is affirmative, then this condition is also sufficient for the existence of a solution, without need for the additional assumption that (M, d) be finitely compact.

THEOREM 7.5. *The Problem for Metric Spaces admits a solution for the finitely compact space* (M, d) *if and only if*

(1) *between each pair of points in M there exists a rectifiable curve, and*

(2) *for each $x \in M$ and each $\epsilon > 0$, there is a $\delta > 0$ such that if $d(x, x') < \delta$ then there is a rectifiable curve in M from x to x' with length less than ϵ.*

Proof left to the reader. A similar result appears in Menger [7].

By analogy with the above Problem for Metric Spaces, we may also formulate a

PROBLEM FOR METRIZABLE SPACES. *Given a metrizable space M, find an intrinsic metric for M which is consistent with the topology on M. Equivalently, given a metric space (M, d), find an intrinsic metric d^* for M which is equivalent to d.*

The following result is due to Moise [5] and Bing [6], independently.

THEOREM 7.6. (Bing, Moise). *Every compact, connected, locally connected metrizable space (that is, every Peano space) admits an intrinsic metric consistent with its topology.*

8. M-convexity. The following definition is due to Menger [7].

DEFINITION 8.1. *A metric space (M, d) is convex in the sense of Menger, or M-convex for short, if for each pair of distinct points x, $z \in M$ there exists a third point $y \in M$ such that $d(x, y) + d(y, z) = d(x, z)$.*

On intuitive grounds, the two properties (1) d is an intrinsic metric, (2) (M, d) is M-convex, seem to be closely related. They are not quite equivalent.

Examples. (1) If d is intrinsic, (M, d) need not be M-convex. For example, let M be the complement in R^2 of an open line segment and d the Euclidean metric.

(2) If (M, d) is M-convex, it need not even be connected. For example, let M be the complement in R^1 of a single point, and d the Euclidean metric.

(3) If (M, d) is M-convex and connected, d need not be intrinsic. Let d be the usual intrinsic metric on a circle, and obtain M by removing a single point from the circle. Restricting d to M, the resulting (M, d) is connected and M-convex, but d is no longer intrinsic on M.

On the positive side, we do have

THEOREM 8.2. *If (M, d) is finitely compact, then it is M-convex if and only if d is intrinsic.*

Proof left to the reader.

The author is an Alfred P. Sloan Research Fellow.

References

1. H. Busemann, The Geometry of Geodesics, Academic Press, New York, 1955.
2. W. Rinow, Die innere Geometrie der metrischen Räume, Springer-Verlag, Berlin, 1961.
3. H. Busemann, Local metric geometry, Trans. Amer. Math. Soc., 56 (1944) 200–274.
4. G. Ascoli, Le curve limite di una varietà data di curve, Atti Accad. Naz. Lincei, Rome, (III) 18 (1883) 551–586.
5. E. E. Moise, Grille decomposition and convexification theorems for compact metric locally connected continua, Bull. Amer. Math. Soc., 55 (1949) 1111–1121.
6. R. H. Bing, Partitioning a set, *ibid.*, 1101–1110.
7. K. Menger, Untersuchungen über allgemeine Metrik, Math. Ann., 100 (1928) 75–163.

HOMAGE TO JOHN RAINWATER

Hom, hom on the range,
Where the co and the contra do play,
Where seldom is heard
A non-natural word,
And the arrows all run the right way.

CHRISTIAN ROSELIUS, Tulane University

THE GAMMA FUNCTION WITH VARYING DIFFERENCE INTERVAL

TOMLINSON FORT, Emory University

Before actually discussing the gamma function it is necessary to introduce a section on *general periodic functions* and one on a generalization of the *Euler-Maclaurin sum formula*.

1. General periodic functions. Let $z=x+yi$ and $h(z)=p(z)+iq(z)$, where x, y, $p(z)$ and $q(z)$ are real. Then $f(z)$ is called general periodic with period $h(z)$, over a domain R, if

$$f(z+h(z)) = f(z), \tag{1}$$

whenever z and $z+h(z)$ belong to R.

An example of a general periodic function is $\sin z^2$ where $h(z) = -z + \sqrt{(z^2+2\pi)}$. Any determination may be given to $\sqrt{(z^2+2\pi)}$. Other examples are easy to construct.

We now place certain restrictions on $h(z)$ and under these restrictions show that general periodic functions with period $h(z)$, other than constants, exist. These restrictions are not necessary. The reader can formulate other sufficient conditions.

Assume $0<\epsilon<p(z)$ and that if z and $\bar{z}$ are different numbers in R then $z+h(z)\neq\bar{z}+h(\bar{z})$. We now draw the lines $x=\epsilon$, $x=2\epsilon, \cdots, x=n\epsilon \cdots$ in the complex plane. These lines divide the right-hand half of the complex plane into strips. We now assign values to $f(z)$ at all points of the strip $0\leqq x\leqq\epsilon$. From these values determine $f(z)+h(z)$ by equation (1). If $f(z)$ is not determined at all points of the strip $\epsilon<x\leqq 2\epsilon$ assign it a value at each of these points as desired. Now determine $f(z)$ for the points of the strip $2\epsilon<x\leqq 3\epsilon$ from its values on the strip $\epsilon<x\leqq 2\epsilon$ by (1) so far as is possible and so far as $f(z)$ has not already been determined by the points of the strip $0\leqq x\leqq\epsilon$. If there exist now points of the strip $2\epsilon<x\leqq 3\epsilon$ where $f(z)$ has not been defined we assign values to $f(z)$ at these points as desired. We now consider the strip $3\epsilon<x\leqq 4\epsilon$. So far as $f(z)$ can not be defined at points of this strip from its values on the strip $0\leqq x<3\epsilon$ we assign $f(z)$ values at these points as desired. We continue to proceed in this manner until $f(z)$ is defined at all points of the half plane $x\geqq 0$. It is immediate that $f(z)$ is general periodic with period $h(z)$ over this half plane. It is in general discontinuous.

2. The Euler-Maclaurin Formula. Let $h(x)$ be real and positive, and let $x+h(x)$ increase monotonically with x. Now let $x_1=x$, $x_2=x_1+h(x_1), \cdots, x_n=x_{n-1}+h(x_{n-1})$. Then let $x_n-x_1=h_n(x)$. Then $h_n(x)=h_{n-1}(x)+h(x_{n-1})$ and $h_n(x)=\sum_{i=1}^{n-1} h(x_i)$. In the work which follows $B_n(x)$ is the Bernoulli polynomial of order n, and $\overline{B}_n(x)$ is the function with period 1 coinciding with $B_n(x)$ over the interval $[0, 1]$. We also let B_n be the Bernoulli member of order n and $Q_n(x)=B_n(x)-B_n$, $\overline{Q}_n(x)=\overline{B}_n(x)-B_n$. We remark that $(d/dx)B_n(x)=nB_{n-1}(x)$. For

other properties of the Bernoulli polynomials and numbers see, for example, [1], Chapter III. We now write down

$$-R_m(a) = h^m(a)\int_0^1 \frac{\overline{B}_m(w-t)}{m!} F^{(m)}(a+h(a)t)\,dt. \tag{2}$$

Carrying out certain operations on this as explained in detail in [1], page 51, we arrive at the formula,

$$\begin{aligned} F(a+wh(a)) &= \frac{1}{h(a)}\int_a^{a+h(a)} F(t)\,dt + \sum_{v=1}^{m} h^{v-1}(a)\,\frac{B_v(w)}{v!}\,\Delta F^{(v-1)}(a) \\ &\quad - h^m(a)\int_0^1 \frac{\overline{B}_m(w-t)}{m!} F^{(m)}(a+h(a)t)\,dt. \end{aligned} \tag{3}$$

Here $\Delta F^{(v-1)}(a) = F^{(v-1)}(a+h(a)) - F^{(v-1)}(a)$. The parenthetical superscripts denote differentiation. Formula (3) is the basic Euler-Maclaurin formula.

Now let $w=0$, $m=2k$. Note that $Q_{2k}(1-t)=Q_{2k}(t)$ and that $B_{2k-1}=0$ when $k>1$. We then let a equal successively $x_1, x_2, \cdots, x_n$ and sum. If $k>1$ we get

$$\begin{aligned} \sum_{i=1}^{n} F(x_i)h(x_i) &= \sum_{i=1}^{n}\int_{x_i}^{x_{i+1}} F(t)\,dt + \sum_{i=1}^{n}\sum_{v=1}^{2k-2}\frac{B_v}{v!}\,\Delta F^{(v-1)}(x_i)h^v(x_i) \\ &\quad - \frac{1}{(2k)!}\int_0^1\left[Q_{2k}(t)\sum_{i=1}^{n} F^{(2k)}(x_i+h(x_i)t)h^{2k+1}(x_i)\right]dt. \end{aligned} \tag{4}$$

Denote the last expression by $-R_{2k}^1$. We have

$$R_{2k}^1 = -\frac{1}{(2k)!}\int_0^1\left[Q_{2k}(t)\sum_{i=1}^{n} F^{(2k)}(x_i+h(x_i)t)h^{2k+1}(x_i)\right]dt. \tag{5}$$

Since $Q_{2k}(t)$ retains the same sign over the interval (0, 1) we can apply the first law of the mean for integrals. We get

$$R_{2k}^1 = \left[-\int_0^1 Q_{2k}(t)\,dt\right]\frac{1}{(2k)!}\sum_{i=1}^{n} F^{(2k)}(x_i+\theta(x,n)h(x_i))h^{2k+1}(x_i),$$

$0<\theta<1$. But $Q_{2k}(t) = (1/(2k+1))Q'_{2k+1}(t) - B_{2k}$. Hence $\int_0^1 Q_{2k}(t)\,dt = -B_{2k}$. Hence

$$R_{2k}^1 = \frac{B_{2k}}{(2k)!}\sum_{i=1}^{n} F^{(2k)}(x_i+\theta(x,n)h(x_i))h^{2k+1}(x_i). \tag{6}$$

Since $B_{2k}/2k!$ is bounded,

$$\left|R_{2k}^1\right| < M\sum_{i=1}^{n}\left|F^{(2k)}(a+\theta(x,n)h(x_i))h^{2k+1}(x_i)\right|.$$

Under certain conditions we can obtain other forms for R_{2k}^1. Let us assume that $F^{(i)}(t)$ retains the same sign when $t>0$ and that $F^{2i}(t)\cdot F^{2i-2}(t)>0$ when $t>0$ and

$j = 2, 3, \cdots$. We note that $Q'_{2k}(t) = 2kQ_{2k-1}(t)$, $k > 1$ and that $Q'_{2k-1}(t) = (2k-1)Q_{2k-2} - B_{2k-2}$ and that $Q_{2k}(t) \cdot Q_{2k-2}(t) < 0$, $0 < t < 1$. We now consider (5) and integrate by parts twice. We obtain

$$\begin{aligned} \overset{1}{R}_{2k} = &- \frac{B_{2k-2}}{(2k-2)!} \sum_{i=1}^{n} \Delta F^{(2k-3)}(x_i) h^{2k-2}(x_i) \\ &- \frac{1}{(2k-2)!} \int_0^1 \left[Q_{2k-2}(t) \sum_{i=1}^{n} F^{(2k-2)}(x_i + h(x_i)t) h^{2k-1}(x_i) \right] dt. \end{aligned} \tag{7}$$

Now if $A = B + C$ and $AC < 0$ then $A = \theta B$, $0 < \theta < 1$. Hence

$$\overset{1}{R}_{2k} = \frac{B_{2k-2}}{(2k-2)!} \theta \sum_{i=1}^{n} (\Delta F^{(2k-3)}(x_i)) h^{2k-2}(x_i). \tag{8}$$

We assume, of course, the existence of all derivatives that enter any formula.

If we advance k by 1 we have

$$\overset{1}{R}_{2k+2} = \frac{B_{2k}}{2k!} \theta \sum_{i=1}^{n} (\Delta F^{(2k-1)}(x_i)) h(x_i)^{2k}. \tag{9}$$

3. The gamma function. We shall solve the difference equation

$$\frac{\Delta u(x)}{h(x)} = \ln x, \qquad x \geqq \epsilon > 0. \tag{10}$$

Here $\Delta u(x) = u(x + h(x)) - u(x)$. We shall require that $0 < c < h(x) < E$, where c and E are constants. We shall also require that $\sum_{i=1}^{\infty} (\ln x_i) \Delta h(x_i)$ converge uniformly $x \geqq \epsilon$. A function that satisfies these conditions is

$$h(x) = 1 + \frac{1}{x}, \qquad x \geqq 1.$$

To show that $\sum_{i=1}^{\infty} (\ln x_i) \Delta h(x_i)$ converges note that $x + i - 1 < x_i < x + i - 1 + (i-1)/x$. Hence,

$$\begin{aligned} \Delta h(x_i) &= \frac{h(x_i)}{x_i(x_i + h(x_i))} < \frac{1 + (1/c)}{(x + i - 1)\left(x + i - 1 + \dfrac{1}{x + i - 1 + (i-1)/x}\right)} \\ &< \frac{1 + (1/c)}{(x + i - 1)^2}. \end{aligned}$$

$$\begin{aligned} \ln x_i &\leqq \ln\left(x + i - 1 + \frac{i-1}{x}\right) = \ln\left(x + (i-1)\left(1 + \frac{1}{x}\right)\right) \\ &\leqq \ln(x + 2(i-1)) \end{aligned}$$

if $x \geqq 1$. Hence,

$$(\ln x_i)\Delta h(x_i) < \left(1+\frac{1}{c}\right)\frac{\ln(x+2(i-1))}{(x+i-1)^2} < \left(1+\frac{1}{c}\right)\frac{2}{(x+i-1)^{3/2}}$$
$$< \left(1+\frac{1}{c}\right)\frac{2}{(i-1)^{3/2}}, \qquad i>1.$$

This is the general term of a convergent series of constants. Consequently $\sum_{i=1}^{\infty}(\ln x_i)\Delta h(x_i)$ converges uniformly, $x\geqq 1$.

We note that

$$\Delta h(x_i) = h(x_i + h(x_i)) - h(x_i) = h(x_{i+1}) - h(x_i).$$

It will be noticed that, although the work is written for $\ln x$, we can replace $\ln x$ by $F(x)$ with only trivial modifications if we require that $F^{(2j)}(x)$ exist and retain the same sign when $x\geqq 1$ and that $F^{(2j)}(x)F^{2j-2}(x)>0$ when $1\leqq j\leqq k$ and that $\sum_{i=1}^{\infty}F(x_i)\Delta h(x_i)$ converge uniformly when $x\geqq\epsilon$ and that $F^{(v)}(x)x^{(v-1)}$ approaches 0 when x becomes infinite, $v=1, 2, \cdots$.

THEOREM. *If* $\sum_{i=1}^{\infty}(\ln x_i)\Delta h(x_i)$ *converges uniformly*, $x\geqq\epsilon>0$, *and if* $0<c<h(x)<E$, *where* c *and* E *are constants, then*

$$\int_1^{x_{n+1}}\ln t\,dt - \sum_{i=1}^{n}(\ln x_i)h(x_i) - \tfrac{1}{2}(\ln x_{n+1})h(x_{n+1}) \tag{11}$$

has a limit as n *becomes infinite and the first difference of this limit is* $(\ln x)h(x)$.

We shall call this limit $\ln G(x)$. We are to prove

$$\frac{\Delta\ln G(x)}{h(x)} = \frac{\ln G(x+h(x)) - \ln G(x)}{h(x)} = \ln x.$$

Proof. We rewrite formula (4), replacing $F(x)$ by $\ln x$ and changing signs throughout the equation. We use formula (6) for the remainder.

$$\begin{aligned}\int_x^{x_{n+1}}\ln t\,dt - \sum_{i=1}^{n}(\ln x_i)h(x_i) &= -\sum_{i=1}^{n}\sum_{v=1}^{2k-2}\frac{B_v}{v!}(\Delta\ln^{(v-1)}x_i)h^v(x_i)\\ &\quad - \frac{B_{2k}}{(2k)!}\sum_{i=1}^{n}\ln^{(2k)}(x_i+\theta h(x_i))h^{2k+1}(x_i).\end{aligned}\tag{12}$$

Consider the last sum which we call R'_{2k}. Now $B_{2k}/(2k)!$ is bounded and

$$\begin{aligned}\ln^{(2k)}(x_i+\theta h(x_i))h^{2k+1}(x_i) &= \frac{(2k-1)!h^{2k+1}(x_i)}{[x_i+\theta h(x_i)]^{2k}}\\ &= (2k-1)!\left[\frac{h(x_i)}{x_i+\theta h(x_i)}\right]^{2k}h(x_i) < N\left[\frac{1}{x_i}\right]^{2k}\end{aligned}$$

since $h(x)$ is bounded. Here N is a constant. Consequently

$$\frac{B_{2k}}{(2k)!}\sum_{i=1}^{\infty}(2k-1)!\left[\frac{h(x_i)}{x_i+\theta h(x_i)}\right]^{2k}h(x_i) < \overline{N}\sum_{i=1}^{\infty}\frac{1}{x_i^{2k}}\cdot$$

We call attention to the inequality $0<c<h(x)<E$, whence

$$\sum_{i=1}^{\infty}\frac{1}{x_i^{2k}} < \sum_{i=1}^{\infty}\frac{1}{[x+(i-1)c]^{2k}}\cdot \tag{13}$$

This series converges uniformly when $2k>1$, which we assume. It follows that when $n\rightarrow\infty$, R'_{2k} approaches a function of x defined by a uniformly convergent series in x when $x\geqq\epsilon>0$. This function multiplied by x^{2k-2} approaches zero when x becomes infinite.

We now consider the sum $-S=S_1+S_2+S_3+\cdots+S_{2k-2}$, where

$$S_1 = \sum_{i=1}^{n} B_1[\ln(x_i+h(x_i)) - \ln x_i]h(x_i)$$

$$S_2 = \sum_{i=1}^{n}\frac{B_2}{2!}\left[\frac{1}{x_{i+1}}-\frac{1}{x_i}\right]h^2(x_i)$$

$$S_3 = \sum_{i=1}^{n}\frac{B_3}{3!}\left[\frac{-1}{x_{i+1}^2}+\frac{1}{x_i^2}\right]h^3(x_i)$$

$$S_4 = \sum_{i=1}^{n}\frac{B_4}{4!}\left[\frac{2!}{x_{i+1}^3}-\frac{2!}{x_i^3}\right]h^4(x_i)$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$S_{2k-2} = \sum_{i=1}^{n}\frac{B_{2k-2}}{(2k-2)!}\left[\frac{(2k-4)!}{x_{i+1}^{2k-3}}-\frac{(2k-4)!}{x_i^{2k-3}}\right]h^{2k-2}(x_i).$$

Now let n become infinite. Each of the rows above becomes an infinite series. Note that $x_{i+1}-x_i=h(x_i)$. Consider

$$S_2 = \sum_{i=1}^{\infty}\frac{B_2}{2!}\left[\frac{1}{\xi}\right]^2 h^3(x_i), \qquad x_i<\xi<x_{i+1}.$$

Since $0<c<h(x)<E$

$$\frac{h^3(x_i)}{\xi^2} < \frac{E^3}{[x+c(i-1)]^2}\cdot$$

Consequently when n becomes infinite S_2 becomes an absolutely uniformly convergent series in x when $x\geqq\epsilon>0$. Clearly $S_3\cdots S_{2k}$ can be treated precisely as S_2 with the same result.

Now consider the first of the above sums, namely

$$S_1 = B_1\sum_{i=1}^{n}[\ln x_{i+1}-\ln x_i]h(x_i) = B_1\sum_{i=1}^{n}(\Delta\ln x_i)h(x_i).$$

Summation by parts yields

$$S_1 = -\tfrac{1}{2}\Big[(\ln x_{n+1})h(x_{n+1}) - (\ln x)h(x) - \sum_{i=1}^{n} (\ln x_{i+1})\Delta h(x_i)\Big].$$

Let n become infinite. We transpose $-\frac{1}{2}(\ln x_{n+1})h(x_{n+1})$ in (12). The infinite series $\sum_{i=1}^{\infty} (\ln x_{i+1})\Delta h(x_i)$ converges since $\sum_{i=1}^{\infty} (\ln x_i)\Delta h(x_i)$ converges by hypothesis, and since, on account of the boundedness of h, $\ln x_{i+1}/\ln x_i$ approaches 1.

We now add $\int_1^x \ln t\, dt$ to both sides of (12). All that remains in the right member of (12) approaches a limit as n becomes infinite. The left member is that given in the theorem. We have

$$\begin{aligned}\int_1^{x_{n+1}} \ln t\, dt - \sum_{i=1}^{n} (\ln x_i)h(x_i) - \tfrac{1}{2} (\ln x_{n+1})h(x_{n+1})& \\ = x\ln x - x + 1 + \tfrac{1}{2} (\ln x)h(x) + \tfrac{1}{2} \sum_{i=1}^{n} (\ln x_{i+1})\Delta h(x_{i+1})& \\ - \sum_{i=1}^{n}\sum_{v=2}^{2k-2} \frac{B_v}{v!} [\Delta \ln^{(v-1)} x_i]h^v(x_i) + R'_{2k},&\end{aligned}$$

where $R'_{2k} = (B_{2k}/2k!) \sum_{i=1}^{n} [\ln^{(2k)} (x_i + \theta h(x_i))]h^{2k+1}(x_i)$.

We now shall show that $\ln G(x)$ satisfies (10). Consider (11) which we treat in three parts.

(a) $$\Delta \int_1^{x_{n+1}} \ln t\, dt = \int_{x_{n+1}}^{x_{n+2}} \ln t\, dt = (x_{n+2} - x_{n+1}) \ln \xi,$$

where $x_{n+1} < \xi < x_{n+2}$,

(b) $$\Delta\Big(-\sum_{i=1}^{n} (\ln x_i)h(x_i)\Big) = -(\ln x_{n+1})h(x_{n+1}) + (\ln x_1)h(x_1)$$

(c) $$\Delta[-\tfrac{1}{2}(\ln x_{n+1})h(x_{n+1})] = -\tfrac{1}{2}[(\ln x_{n+1})\Delta h(x_{n+1}) + h(x_{n+2})\Delta \ln x_{n+1}].$$

We now consider these three results. From (a) and (b), since $h(x_{n+1}) = x_{n+2} - x_{n+1}$,

$$(\ln \xi - \ln x_{n+1})h(x_{n+1}) = \frac{1}{\eta} (\xi - x_{n+1})h(x_{n+1}),$$

where $x_{n+1} < \eta < \xi$. Since $h(x)$ is bounded and $\xi - x_{n+1} < h(x_{n+1})$, this approaches zero. Moreover, from (c), $(\ln x_{n+1})\Delta h(x_{n+1})$ approaches zero because it is the general term of a convergent series. Also $h(x_{n+2})\Delta \ln(x_{n+1})$ approaches zero since $h(x)$ is bounded and

$$\Delta(\ln x_{n+1}) = (x_{n+2} - x_{n+1}) \frac{1}{\xi} = h(x_{n+1}) \frac{1}{\xi}\cdot$$

We now total what we have obtained from (a), (b) and (c). All that we get is $\ln x_1 h(x_1)$. But $x_1 = x$. Hence the theorem is proved.

4. Asymptotic form. If we refer to formula (12) with the addition of $\int_1^x \ln t\, dt$ to both sides and the transposition of $\frac{1}{2}(\ln x_{n+1})h(x_{n+1})$ to the left member, and then let n become infinite, we have

$$\ln G(x) = x \ln x - x + 1 + \tfrac{1}{2}(\ln x)h(x) + \tfrac{1}{2}\sum_{i=1}^{\infty}(\ln x_{i+1})\Delta h(x_i) \tag{14}$$
$$- \sum_{i=1}^{\infty}\sum_{v=2}^{2k-2}\frac{B_v}{v!}\left[\Delta \ln^{(v-1)}(x_i)\right]h^v(x_i) + R'_{2k},$$

where

$$R'_{2k} = \frac{B_{2k}}{2k!}\sum_{i=1}^{\infty}\left[\ln^{(2k)}(x_i + \theta h(x_i))\right]h^{2k+1}(x_i). \tag{15}$$

Alternate forms for the remainder can be found. If, for example, we refer to formula (8) we have

$$R'_{2k} = -\theta\frac{B_{2k}}{2k!}\sum_{i=1}^{\infty}(\Delta \ln^{2k-1} x_i)h^{2k}(x_i).$$

Now if we let $h(x_i) \equiv 1$ in (14) and perform the differentiation on $\ln x$ we get the following form:

$$\ln G(x) = (x + \tfrac{1}{2})\ln x - x + 1 - \sum_{v=2}^{2k-1}\frac{B_v}{v(v-1)}\cdot\frac{1}{x^{v-1}} + R_{2k}. \tag{16}$$

If we use the second formula for R_{2k} we have

$$R'_{2k} = \theta\frac{B_{2k}}{2k(2k-1)}\cdot\frac{1}{x^{2k-1}}\,. \tag{17}$$

This differs from the classical formula for $\ln \Gamma(x)$ (see [1] p. 61) only by the absence of $\ln\sqrt{2\pi}$ and the presence of 1. However, any general periodic function can be added to a solution of (10) and the result will still be a solution. We consequently add $\ln\sqrt{2\pi} - 1$ to the right member of formula (14). We denote the function that we obtain by $\ln \Gamma_h(x)$. We have

$$\ln \Gamma_h(x) = \ln\sqrt{2\pi} + x\ln x - x + \tfrac{1}{2}(\ln x)h(x) + \tfrac{1}{2}\sum_{i=1}^{\infty}(\ln x_{i+1})\Delta h(x_i)$$
$$+ \sum_{i=1}^{\infty}\sum_{v=2}^{2k-2}\frac{B_v}{v!}\left[(\Delta \ln^{(v-1)}x_i)h^v(x_i)\right] + R'_{2k},$$

where R'_{2k} is given by (15) or (17).

We note particularly from (17) that $x^{2k-2}R'_{2k}$ approaches zero when x becomes infinite. Since $\Delta \ln \Gamma(x) = h(x)\ln x$ then $\Gamma_h(x + h(x)) = x^{h(x)}\Gamma_h(x)$.

5. The gamma function in the complex plane. Draw a line parallel to the axis of imaginaries and distant $\epsilon>0$ from it. All variables are confined to the half-plane $x\geqq\epsilon$. We assume $0<\eta<|h(z)|<E$ also $\mathrm{Re}(z+h(z))>0$. We require that $h(z)$ be real when $z=x$ is real. In addition we assume that $h(z)$ is analytic over the half-plane in question. We also assume that $h(x)$ meets all the requirements previously put upon it. We let $\ln z=\ln\sqrt{x^2+y^2}+\phi i$, where $-\pi/2<\phi<\pi/2$. The points z_i are determined by the equations $z_1=z$, $z_i=z_{i-1}+h(z_{i-1})$, $i>1$.

We write down the expression

$$\begin{aligned}\ln\sqrt{2\pi}+z\ln z-z+\tfrac{1}{2}(\ln z)h(z)+\tfrac{1}{2}\sum_{i=1}^{\infty}(\ln z_{i+1})\Delta h(z_i)\\ +\sum_{i=1}^{\infty}\sum_{v=2}^{2k-2}\frac{B_v}{v!}[\Delta\ln^{(v-1)}(z_i)]h^v(z_i)+R'_{2k},\end{aligned}\tag{18}$$

where

$$R'_{2k}=\frac{1}{2k!}\int_0^1 Q_{2k}(t)\left[\sum_{i=1}^{\infty}\ln^{(2k)}(z_i+h(z_i)t)h^{2k+1}(z_i)\right]dt.$$

All series appearing here converge uniformly over the half-plane $x\geqq\epsilon$, the first by assumption and the others by easy proof (see par. 3). Formula (18) consequently defines an analytic function. This function reduces to $\ln\Gamma_h(x)$ when z is real. This is the same function for different values of k. Suppose that there were two. These two would both be analytic and each would reduce to $\Gamma_h(x)$ when z is real. They are consequently identical from the general theory of analytic functions. We denote the function defined by (18) by $\ln\Gamma_h(z)$. Similarly the relation $\Delta\ln\Gamma_h(z)/h(z)=\ln z$ will hold for complex z from the general theory of analytic functions.

6. Another generalization. Let us consider the equation

$$\Delta u(x)=\ln x.\tag{19}$$

We write this $\Delta u(x)/h(x)=\ln(x)/h(x)$. Replacing $\ln x/h(x)$ by $F(x)$ we have

$$\frac{\Delta u(x)}{h(x)}=F(x).$$

Now if $h(x)$ is as previously and $F(x)$ is such that $F^{(2x)}$ retains the same sign and $F^{(2k)}(x)F^{(2k-2)}(x)>0$ and $F^{(2k)}(x)\cdot x^{2k-2}$ approaches zero when x becomes infinite, then (19) can be solved as (10) was. The solution which reduces to $\Gamma(x)$ when $h(x)=1$ is a generalization of $\Gamma(x)$. The requirements on F place further restrictions on $h(x)$. If, for example, $h(x)=\ln x/(x\ln x-x)$, $x>e^2$ then all requirements are fulfilled. We call this generalization $\Gamma^h(x)$. Clearly

$$\Gamma^h(x+h(x))=x\Gamma^h(x).\tag{20}$$

This paper is part of a general project on which the author is working with the support of The National Aeronautics and Space Administration.

References

1. T. Fort, Finite Differences and Difference Equations in the Real Domain, Clarendon Press, Oxford, 1948.
2. N. E. Nörlund, Differenzenrechnung, Springer, Berlin, 1924.

LOGARITHMIC NUMBERS AND THE FUNCTIONS $d(n)$ AND $\sigma(n)$

J. M. GANDHI, Univ. of Rajasthan, Jaipur (India)

1. In this paper we derive some formulae relating logarithmic numbers $G_\gamma^{(n)}(-1)$ and $G_\gamma^{(n)}(1)$, $G_\gamma^{(n)}(t)$ being defined by the generating function, [2]

$$e^{-xt}\log(1-x^n) = -\sum_{\gamma=1}^{\infty} \frac{G_\gamma^{(n)}(t)x^\gamma}{\gamma!} \tag{1}$$

with the function $d(n)$, the number of divisors of n, $\sigma(n)$, the sum of the divisors of n, $S(n)$, the odd divisors of n, $C(2n)$, sum of even divisors of $2n$ and $\alpha(n)$, excess of the sum of odd divisors of n over the sum of even divisors of n.

As corollaries to those relations, the following interesting congruences have been obtained, in which p denotes an odd prime:

$$D_p = \sum_{n=1}^{p} G_p^{(n)}(1) \equiv -1 \pmod{p} \tag{2}$$

$$F_p = \sum_{n=1}^{p} G_p^{(n)}(-1) \equiv -1 \pmod{p} \tag{3}$$

$$H_p = \sum_{n=1}^{p} G_p^{(n)}(1)/n \equiv -2 \pmod{p} \tag{4}$$

$$J_p = \sum_{n=1}^{p} G_p^{(n)}(-1)/n \equiv -2 \pmod{p}. \tag{5}$$

Putting $t=+1$ and -1 in (1) respectively, we get

$$e^{-x}\log(1-x^n) = -\sum_{\gamma=1}^{\infty} \frac{G_\gamma^{(n)}x^\gamma}{\gamma!} \tag{6}$$

and

$$e^{+x}\log(1-x^n) = -\sum_{\gamma=1}^{\infty} \frac{G_\gamma^{(n)}(-1)x^\gamma}{\gamma!}\,. \tag{7}$$

$G_\gamma^{(n)}(1)$ has been abbreviated as $G_\gamma^{(n)}$. In equation (6), putting $n=1, 2, 3, \cdots$, and adding we get

$$e^{-x} \log \left[(1-x)(1-x^2)(1-x^3) \cdots\right] = -\sum_{\gamma=1}^{\infty} \frac{D_\gamma x^\gamma}{\gamma!}, \tag{8}$$

where $D_\gamma = \sum_{n=1}^{\infty} G_\gamma^{(n)} = \sum_{n=1}^{\gamma} G_\gamma^{(n)}$, since $G_\gamma^{(n)}$ for $\gamma > n$ is 0. Now

$$\log \left[(1-x)(1-x^2)(1-x^3) \cdots\right] = -\sum_{\gamma=1}^{\infty} \frac{\sigma(\gamma) x^\gamma}{\gamma}. \tag{9}$$

When (9) is used, (8) becomes

$$e^{-x} \sum_{\gamma=1}^{\infty} \frac{\sigma(\gamma) x^\gamma}{\gamma} = \sum_{\gamma=1}^{\infty} \frac{D_\gamma x^\gamma}{\gamma}. \tag{10}$$

Expanding and equating the coefficients of x^γ we get

$$D_\gamma = \gamma! \sum_{k=0}^{\gamma-1} \frac{(-1)^k \sigma(\gamma-k)}{k!(\gamma-k)}. \tag{11}$$

We can also write (10) as

$$\sum_{\gamma=1}^{\infty} \frac{\sigma(\gamma) x^\gamma}{\gamma} = e^x \sum_{\gamma=1}^{\infty} \frac{D_\gamma x^\gamma}{\gamma!}. \tag{12}$$

Expanding and equating the coefficients of x^γ and rearranging, we get

$$\sigma(\gamma) = \frac{1}{(\gamma-1)!} \sum_{k=0}^{\gamma-1} \binom{\gamma}{k} D_{\gamma-k}. \tag{13}$$

Similarly starting from (7), and putting $F_\gamma = \sum_{n=1}^{\gamma} G_\gamma^{(n)}(-1)$, we can obtain

$$e^x \sum_{\gamma=1}^{\infty} \frac{\sigma(\gamma) x^\gamma}{\gamma} = \sum_{\gamma=1}^{\infty} \frac{F_\gamma x^\gamma}{\gamma!}. \tag{14}$$

Then proceeding as before, we can prove that

$$F_\gamma = \gamma! \sum_{k=0}^{\gamma-1} \frac{\sigma(\gamma-k)}{k!(\gamma-k)}, \tag{15}$$

$$\sigma(\gamma) = \frac{1}{(\gamma-1)!} \sum_{k=0}^{\gamma-1} \binom{\gamma}{k} (-1)^k F_{\gamma-k}. \tag{16}$$

Now multiplying (14) and (10), expanding and equating the coefficients of x^γ and rearranging, we get

$$\sum_{k=1}^{\gamma-1} \frac{\sigma(\gamma-k)\sigma(k)}{(\gamma-k)k} = \frac{1}{\gamma!} \sum_{k=1}^{\gamma-1} \binom{\gamma}{k} D_{\gamma-k} F_k. \tag{17}$$

Also dividing (12) by (14), we get

$$e^{-2x} \sum_{\gamma=1}^{\infty} \frac{F_\gamma x^\gamma}{\gamma!} = \sum_{\gamma=1}^{\infty} \frac{D_\gamma x^\gamma}{\gamma!}, \tag{18}$$

whence we can prove that

$$F_\gamma = \sum_{k=0}^{\gamma-1} D_{\gamma-k} \cdot 2^k \tag{19}$$

and

$$D_\gamma = \sum_{k=0}^{\gamma-1} F_{\gamma-k}(-2)^k. \tag{20}$$

From (19) or (20) we have

$$F_\gamma \equiv D_\gamma \pmod 2. \tag{21}$$

Now from (6), we can easily prove that

$$e^{-x} \log \left[(1-x)(1-x^2)^{1/2}(1-x^3)^{1/3} \cdots \right] = -\sum_{\gamma=1}^{\infty} \frac{H_\gamma x^\gamma}{\gamma!}, \tag{22}$$

where

$$H_\gamma = \sum_{n=1}^{\infty} \frac{G_\gamma^{(n)}}{n} = \sum_{n=1}^{\gamma} G_\gamma^{(n)}/n. \tag{23}$$

Since

$$\log \left[(1-x)(1-x^2)^{1/2}(1-x^3)^{1/3} \cdots \right] = -\sum_{n=1}^{\gamma} \frac{d(\gamma) x^\gamma}{\gamma}, \tag{24}$$

(22) becomes

$$e^{-x} \sum_{n=1}^{\infty} \frac{d(n) x^n}{n} = \sum_{n=1}^{\infty} \frac{H_n x^n}{n}. \tag{25}$$

Similarly, from (7) we can prove that

$$e^{x} \sum_{n=1}^{\infty} \frac{d(n) x^n}{n} = \sum_{n=1}^{\infty} \frac{J_n x^n}{n}, \tag{26}$$

where

$$J_\gamma = \sum_{n=1}^{\gamma} \frac{G_\gamma^{(n)}(-1)}{n}. \tag{27}$$

Proceeding as before, we can show that

$$H_\gamma = \gamma! \sum_{k=0}^{\gamma-1} \frac{d(\gamma-k)(-1)^k}{k!(\gamma-k)} \tag{28}$$

$$d(\gamma) = \frac{1}{(\gamma-1)!} \sum_{k=0}^{\gamma-1} \binom{\gamma}{k} H_{\gamma-k} \tag{29}$$

$$J_\gamma = \gamma! \sum_{k=0}^{\gamma-1} \frac{d(\gamma-k)}{k!(\gamma-k)} \tag{30}$$

$$d(\gamma) = \frac{1}{(\gamma-1)!} \sum_{k=0}^{\gamma-1} (-1)^k \binom{\gamma}{k} J_{\gamma-k} \tag{31}$$

$$\sum_{k=1}^{\gamma-1} \frac{d(\gamma-k)d(k)}{(\gamma-k)k} = \frac{1}{\gamma!} \sum_{k=1}^{\gamma-1} \binom{\gamma}{k} H_k J_{\gamma-k} \tag{32}$$

$$J_\gamma = \sum_{k=0}^{\gamma-1} H_{\gamma-k}(-2)^k \tag{33}$$

$$H_\gamma = \sum_{k=0}^{\gamma-1} J_{\gamma-k}(2)^k. \tag{34}$$

From (33) or (34) it follows that

$$J_\gamma \equiv H_\gamma \pmod 2). \tag{35}$$

Now putting $n=1, 3, 5, \cdots$, in (7) and (8), and adding the expressions so obtained, we get

$$e^{-x} \log [(1-x)(1-x^3)(1-x^5) \cdots] = -\sum_{\gamma=1}^{\infty} \sum_{n=1}^{\gamma}{}^* \frac{G_\gamma^{(n)} x^\gamma}{\gamma!} \tag{36}$$

$$e^{x} \log [(1-x)(1-x^3)(1-x^5) \cdots] = -\sum_{\gamma=1}^{\infty} \sum_{n=1}^{\gamma}{}^* \frac{G_\gamma^{(n)}(-1)x^\gamma}{\gamma!}, \tag{37}$$

where $\sum^*$ denotes that the summation is taken over only the odd values of $n \leqq \gamma$.

We have

$$\log [(1-x)(1-x^3)(1-x^5) \cdots] = -\sum_{\gamma=1}^{\infty} \frac{S(\gamma)x^\gamma}{\gamma}, \tag{38}$$

$S(\gamma)$ being already defined.

Using (36), (37) and (38), we can derive results similar to equations (28) to (34).

Proceeding as before, we can prove the following set of seven general formulae

$$\frac{\phi(\gamma)}{\gamma!}=\sum_{k=0}^{\gamma-1}\frac{\theta(\gamma-k)(-1)^k}{k!(\gamma-k)} \tag{39}$$

$$\theta(\gamma)=\frac{1}{(\gamma-1)!}\sum_{k=0}^{\gamma-1}\binom{\gamma}{k}\phi'_{\gamma-k}(-1)^k \tag{40}$$

$$\frac{\phi'(\gamma)}{\gamma!}=\sum_{k=0}^{\gamma-1}\frac{\theta(\gamma-k)}{k!(\gamma-k)} \tag{41}$$

$$\theta(\gamma)=\frac{1}{(\gamma-1)!}\sum_{k=0}^{\gamma-1}\binom{\gamma}{k}\phi_{\gamma-k} \tag{42}$$

$$\sum_{k=1}^{\gamma-1}\frac{\theta(\gamma-k)\theta(k)}{(\gamma-k)k}=\frac{1}{\gamma!}\sum_{k=0}^{\gamma-1}\phi_{\gamma-k}\phi_k'\binom{\gamma}{k} \tag{43}$$

$$\phi'(\gamma)=\sum_{k=0}^{\gamma-1}\binom{\gamma}{k}\phi_{\gamma-k}2^k \tag{44}$$

$$\phi(\gamma)=\sum_{k=0}^{\gamma-1}\binom{\gamma}{k}\phi'_{\gamma-k}(-2)^k. \tag{45}$$

Here θ, ϕ and ϕ' are given in the following table.

TABLE 1

No.	θ_γ	ϕ_γ	ϕ_γ'
1	$\sigma(\gamma)$	D_γ	F_γ
2	$d(\gamma)$	H_γ	J_γ
3	$s(\gamma)$	$\sum_{n=1}^{\gamma}{}^{*}G_\gamma^{(n)}$	$\sum_{n=1}^{\gamma}{}^{*}G_\gamma^{(n)}(-1)$
4	$c(2\gamma)$	$\sum_{n=1}^{2\gamma}{}^{**}G_\gamma^{(n)}$	$\sum_{n=1}^{2\gamma}{}^{**}G_\gamma^{(n)}(-1)$
5	$\alpha(\gamma)$	$\sum_{n=1}^{\gamma}(-1)^{n+1}G_\gamma^{(n)}$	$\sum_{n=1}^{\gamma}(-1)^{n+1}G_\gamma^{(n)}(-1)$

$\sum^*$ has already been defined while $\sum^{**}$ mean that the summation is taken over only the even values of n.

As corollaries to these results, we now prove the congruences (2)–(5). In (13), let $\gamma=p$, p being an odd prime, then since $\sigma(p)=p+1\equiv 1 \pmod p$, $(p-1)!\equiv -1 \pmod p$ and $\binom{p}{k}\equiv 0 \pmod p$ except when $k=0$ or $k=p$, and hence we get

$$D_p\equiv -1 \pmod p,$$

whereby congruence (2) has been proved. Similarly, from (16) congruence (3)

can be proved. Then in (29), putting $\gamma = p$, and since $d(p) = 2$, we get

$$H_p \equiv -2 \pmod{p},$$

whereby congruence (4) has been proved. Similarly, from (31) congruence (5) can be proved.

Acknowledgements: The author is thankful to Prof. H. Gupta and Prof. M. F. Soonawala for their kind encouragement.

References

1. L. E. Dickson, Theory of Numbers, vol. 1, Chelsea, New York, 1952.
2. J. M. Gandhi, On logarithmic numbers, Math. Student, 31(1963) No. 1–2, pp. 73–83.

A COMPARISON OF METRIZABLE AND COMPACT TOPOLOGICAL SPACES

JANET S. ALLSBROOK, University of Florida
(Now at Newberry College, Newberry, S. C.)

It is possible to distinguish a class $\mathcal{C}$ of topological spaces $\{X\}$ with continuous functions on $X \times X$ into $[0, 1]$ and equicontinuous families of functions on X into $[0, 1]$ for $X \in \mathcal{C}$. This class $\mathcal{C}$ of topological spaces can be characterized without the use of the concept of "uniformities" on $X \times X$, $X \in \mathcal{C}$. The class $\mathcal{C}$ is shown, however, to coincide with the class of topological spaces which are uniformizable. An interesting comparison results between two subclasses of $\mathcal{C}$, the metrizable topological spaces and the compact topological spaces, which is described entirely in topological terms. Uniformity properties are used to advantage, however, in deriving the results.

All notation and definitions are consistent with [1] unless stated otherwise. All topological spaces are assumed to be Hausdorff.

1. Metrizable topological spaces. We first characterize topological spaces, $\{X\}$, which are metrizable with a single continuous function on $X \times X$ to $[0, 1]$, $X \in \{X\}$.

DEFINITION 1.1, [1]. *F, a family of functions on $(X, \mathfrak{I})$, a topological space, to $(Y, \mathfrak{U})$, a uniform space, is equicontinuous at $x \in X$ if and only if for each $\alpha \in \mathfrak{U}$, there exists a neighborhood of x, $N(x)$, such that*

(1) *$f(N(x)) \subset \alpha[f(x)]$ for each $f \in F$; i.e.,*

(2) *$\bigcap_{f \in F} f^{-1}(\alpha[f(x)])$ is a neighborhood of x.*

F is an equicontinuous family of functions if it is equicontinuous at each $x \in X$.

DEFINITION 1.2. *Let F be a family of functions on X, a set, to $(Y, \mathfrak{U})$, a uniform space. Let $\mathfrak{A}$ be an open base of $\mathfrak{U}$. Consider*

$$B_F = \left\{ \bigcap_{f \in F} f^{-1}(\alpha[f(x)]) : x \in X, \alpha \in \mathfrak{A} \right\}.$$

Then B_F defines a topology on $(X, \mathfrak{J}_F)$ [1, p. 56, B]. If $\{\bigcap_{f\in F} f^{-1}(\alpha[f(x)]: \alpha \in \mathfrak{A}\}$ is an open neighborhood base for each $x \in X$, then $\mathfrak{J}_F$ is called the least topology on X such that F is equicontinuous.

REMARK. Actually $\mathfrak{J}_F$ is the least topology on X such that F is equicontinuous and which preserves open neighborhood bases in this sense: if $\{\alpha[f(x)]: \alpha \in \mathfrak{A}\}$ is an open neighborhood base for $f(x)$, then $\{\bigcap_{f\in F} f^{-1}(\alpha[f(x)]): \alpha \in \mathfrak{A}\}$ is an open neighborhood base for x. In this case $\mathfrak{J}_F$ is independent of the choice of $\mathfrak{A}$, an open base of $\mathfrak{U}$.

Notation. Let f be a function on $X \times Y$ to Z, where X, Y, Z are sets. Then, for each $x \in X$, f_x will denote the function on Y to Z where $f_x(y) = f(x, y)$ for each $y \in Y$. Also, for each $y \in Y$, f^y will denote the function on X to Z where $f^y(x) = f(x, y)$ for each $x \in X$.

DEFINITION 1.3. *Let X, Y be topological spaces, let Z be a uniform space, and let f be a function on $X \times Y$ to Z. Then, f is bi-equicontinuous if and only if $\{f_x: x \in X\}$ and $\{f^y: y \in Y\}$ are equicontinuous families of functions.*

Notation. Let (X, d) be a metric space. Then, for each $x \in X$, $\epsilon > 0$, $S_\epsilon^d(x)$ will denote the set $\{y: d(x, y) < \epsilon\}$. Let δ denote the usual metric on the real line; i.e., $\delta(x, y) = |x - y|$ for x, y real numbers. Thus $S_\epsilon^\delta(x) = \{y: |x-y| < \epsilon\}$ for x, y real, $\epsilon > 0$.

LEMMA 1.1. *Let (X, d) be a metric space. Then,*

$$S_\epsilon^d(x) = \bigcap_{y \in X} d_y^{-1}(S_\epsilon^\delta(d_y(x))), \qquad x \in X, \epsilon > 0.$$

Proof. $z \in S_\epsilon^d(x)$ if and only if $d(x, z) < \epsilon$ if and only if $|d(x, y) - d(y, z)| < \epsilon$ for each $y \in X$, if and only if $d_y(z) \in S_\epsilon^\delta(d_y(x))$ for each $y \in X$ if and only if $z \in d_y^{-1}(S_\epsilon^\delta(d_y(x)))$ for each $y \in Y$ if and only if $z \in \bigcap_{y\in X} d_y^{-1}(S_\epsilon^\delta(d_y(x)))$.

THEOREM 1.1. *A topological space, $(X, \mathfrak{J})$, is metrizable if and only if there exists one family of functions, F, on X to $[0, 1]$ such that $\mathfrak{J}$ is the least topology making F equicontinuous.*

Proof. First, assume $(X, \mathfrak{J})$ is metrizable. Let d be a metric on $X \times X$ into $[0, 1]$ which has $\mathfrak{J}$ as the corresponding metric topology. For each $\epsilon > 0$ let $\alpha_\epsilon = \{(t, s): |t - s| < \epsilon\}$ for $s, t \in [0, 1]$. Then, $\{\alpha_\epsilon: \epsilon > 0\}$ is an open base for the usual uniformity on $[0, 1]$. Also, $\alpha_\epsilon[t] = S_\epsilon^\delta(t)$, t real. By Lemma 1.1, $S_\epsilon^d(x) = \bigcap_{y\in X} d_y^{-1}(S_\epsilon^\delta(d_y(x)))$ for each $x \in X$, $\epsilon > 0$. Since, for each $x \in X$, $\{S_\epsilon^d(x): \epsilon > 0\}$ is an open neighborhood base, it follows that $\mathfrak{J}$ is the least topology on X making $\{d_x: x \in X\}$ equicontinuous.

Secondly, let F be a family of functions on $(X, \mathfrak{J})$, a topological space to $[0, 1]$ such that $\mathfrak{J}$ is the least topology making F equicontinuous; i.e., for each

$x\in X$, $\{\bigcap_{f\in F} f^{-1}(S_\epsilon^\delta(f(x)),\ \epsilon>0\}$ is an open neighborhood base. We shall show that $(X, \mathfrak{I})$ is metrizable.

Let $d(x, y) = \text{l.u.b.}_{f\in F}\ \{|f(x)-f(y)|\}$. Then, d is a pseudometric on X; i.e., $d(x, x)=0$ for $x\in X$, $d(x, y)=d(y, x)$ for $x, y\in X$, and $d(x, z)\leqq d(x, y)+d(y, z)$ for $x, y, z\in X$.

Let $\mathfrak{I}_d$ denote the pseudo-metric topology. Then, $\mathfrak{I}_d\subset\mathfrak{I}$. Let $\epsilon>0$, $x\in X$. It is sufficient to find a neighborhood of x, $N_{\mathfrak{I}}(x)$, relative to the $\mathfrak{I}$ topology on X such that $N_{\mathfrak{I}}(x)\subset S_\epsilon^d(x)$. F is equicontinuous on $(X, \mathfrak{I})$ to $[0, 1]$. Therefore, there exists a $\mathfrak{I}$-neighborhood of x, $N_{\mathfrak{I}}^1(x)$, such that $f(N_{\mathfrak{I}}^1(x))\subset S_{\epsilon/2}^\delta(f(x))$ for each $f\in F$. Thus, $y\in N_{\mathfrak{I}}^1(x)$ implies that $|f(y)-f(x)|<\epsilon/2$ for each $f\in F$, and therefore the $\text{l.u.b.}_{f\in F}\ \{|f(x)-f(y)|\}\leqq\epsilon/2<\epsilon$. Hence, $y\in N_{\mathfrak{I}}^1(x)$ implies that $d(x, y)<\epsilon$ or $y\in S_\epsilon^d(x)$. We let $N_{\mathfrak{I}}(x)=N_{\mathfrak{I}}^1(x)$.

Finally, we show that $\mathfrak{I}\subset\mathfrak{I}_d$. Since $\mathfrak{I}$ is the least topology on X making F equicontinuous, for each $x\in X\{\bigcap_{f\in F} f^{-1}(S_\epsilon^\delta(f(x)):\ \epsilon>0\}$ is an open neighborhood base. It is sufficient to show that $S_\epsilon^d(x)\subset\bigcap_{f\in F} f^{-1}(S_\epsilon^\delta(f(x)))$ for each $x\in X$, $\epsilon>0$. Let $y\in S_\epsilon^d(x)$. Then, $d(x, y)<\epsilon$, and, hence, $\text{l.u.b.}_{f\in F}\ \{|f(x)-f(y)|\}<\epsilon$. Hence, $|f(x)-f(y)|<\epsilon$ for each $f\in F$ or $f(y)\in S_\epsilon^\delta(f(x))$ for each $f\in F$ and $y\in\bigcap_{f\in F} f^{-1}(S_\epsilon^\delta(f(x)))$.

Thus, $\mathfrak{I}=\mathfrak{I}_d$. Since $\mathfrak{I}$ is T_1 and pseudo metrizable, it is metrizable.

COROLLARY 1.1. *A topological space, $(X, \mathfrak{I})$ is metrizable if and only if there exists one continuous function f on $X\times X$ to $[0, 1]$ such that $\mathfrak{I}$ is the least topology making f bi-equicontinuous.*

Proof. If $(X, \mathfrak{I})$ is metrizable with metric d on $X\times X$ to $[0, 1]$, d is continuous. Also, $\{d_x: x\in X\}=\{d^x: x\in X\}$. Thus, since $\mathfrak{I}$ is the least topology making $\{d_x: x\in X\}$ equicontinuous, by proof of Theorem 1.1, it is, also, the least making d bi-equicontinuous.

2. Uniformizable topological spaces. We now characterize topological spaces, $\{X\}$, which are uniformizable with a family of continuous functions on $X\times X$ to $[0, 1]$ where $X\in\{X\}$. This characterization of uniform spaces will be used in the next section to establish a result on compact spaces where we take advantage of the fact that such spaces are uniformizable.

THEOREM 2.1. *$(X, \mathfrak{I})$ is uniformizable if and only if there exist some families of functions, i.e., a set $\{F_i: i\in I\}$ of families of functions, on X into $[0, 1]$ such that $\mathfrak{I}$ is the least topology making each F_i equicontinuous for $i\in I$, (a subbase for $\mathfrak{I}$ is $\{\mathfrak{I}_i: i\in I\}$ where for each $i\in I$, $\mathfrak{I}_i$ is the least topology making F_i equicontinuous in the sense of Definition 1.2).*

Proof. First, assume there exists a collection of families of functions, $\{F_i: i\in I\}$ on X to $[0, 1]$ such that $\mathfrak{I}$ is the least topology making each F_i equicontinuous, $i\in I$. Let $\mathfrak{I}_i$ denote the least topology on X such that F_i is equicontinuous, for $i\in I$. Then, by the proof of Theorem 1.1, $(X, \mathfrak{I}_i)$ is pseudometriz-

able. Let d_i be a pseudometric on $X\times X$ into $[0, 1]$ which has $\mathfrak{I}_i$ as its pseudometric topology. Now since $\{S_\epsilon^{d_i}(x): \epsilon>0, x\in X\}$ is a base for $\mathfrak{I}_i$, it follows that a subbase for $\mathfrak{I}$ is $\{S_\epsilon^{d_i}(x): \epsilon>0, x\in X, i\in I\}$. It is easy to see that this latter set, $\{S_\epsilon^{d_i}(x): \epsilon>0, x\in X, i\in I\}$, is, also, a subbase for a uniformizable topology. We let $V_{d_{i,\epsilon}}=\{(x, y): d_i(x, y)<\epsilon\}$. Then, $\{V_{d_{i,\epsilon}}: i\in I, \epsilon>0\}$ is a subbase for a uniformity on $X\times X$, $\mathfrak{U}$ [see 1. pp. 187–8]. Also, $\{V_{d_{i,\epsilon}}[x]: x\in X, \epsilon>0, i\in I\}$ is a subbase for the uniform topology of $\mathfrak{U}$, $\mathfrak{I}\mathfrak{U}$, [1, p. 179, Theorem 5]. But $V_{d_{i,\epsilon}}[x] = S_\epsilon^{d_i}(x)$ for each $x\in X$, $\epsilon>0$, $i\in I$. Thus $\mathfrak{I}=\mathfrak{I}\mathfrak{U}$, and, hence, $\mathfrak{I}$ is uniformizable.

Now, assume $(X, \mathfrak{I})$ is uniformizable with uniformity, $\mathfrak{U}$. Let $\{d_i: i\in I\}$ be the gage for $\mathfrak{U}$. Let $\mathfrak{I}_i$ denote the pseudometric topology of d_i, $i\in I$. Now, $\mathfrak{I}_i$ is the least topology such that $\{(d_i)_x: x\in X\}$ is equicontinuous, by the proof of Theorem 1.1. It is sufficient to show that $\mathfrak{I}$ has $\{\mathfrak{I}_i: i\in I\}$ as a subbase. Now, $\mathfrak{U}$ has $\{V_{d_{i,\epsilon}}: i\in I, \epsilon>0\}$ as a subbase where, as before, $V_{d_{i,\epsilon}}=\{(x, y): d_i(x, y)<\epsilon\}$. The uniform topology, then, has as a subbase, $\{V_{d_{i,\epsilon}}[x]: i\in I, \epsilon>0, x\in X\}$ or equivalently, $\{S_\epsilon^{d_i}(x): i\in I, \epsilon>0, x\in X\}$, [1, p. 187; p. 179, Theorem 5]. Since $\{S_\epsilon^{d_i}(x): \epsilon>0, x\in X\}$ is a base for $\mathfrak{I}_i$, it follows that $\mathfrak{I}$ has $\{\mathfrak{I}_i: i\in I\}$ as a subbase.

COROLLARY 2.1. *A topological space, $(X, \mathfrak{I})$ is uniformizable if and only if there exist some continuous functions, F, on $X\times X$ into $[0, 1]$ such that $\mathfrak{I}$ is the least topology such that each $f\in F$ is biequicontinuous; i.e., it is the least topology making $\{f_x\}$ $[\{f^x\}]$ equicontinuous for each $f\in F$.*

Proof. In the second part of the proof of Theorem 2.1, we let $\{d_i: i\in I\}$ denote the gage of $\mathfrak{U}$, a uniformity for $\mathfrak{I}$, and we denote by $\mathfrak{I}_i$, the pseudometric topology of d_i. It was shown that $\mathfrak{I}$ has a subbase $\{\mathfrak{I}_i: i\in I\}$, where $\mathfrak{I}_i$ is the least topology on X such that $\{(d_i)_x: x\in X\}$ is equicontinuous. But $\{(d_i)_x: x\in X\} =\{(d_i)^x: x\in X\}$. Also, d_i is continuous for each $i\in I$. The necessity of the Corollary follows.

3. Compact topological spaces. We now give a property of compact spaces which shows them to be of the type termed "maximal bi-equicontinuous" while metrizable spaces are equivalent to the type termed "minimal bi-equicontinuous."

DEFINITION 3.1. *Let $(X, \mathfrak{I})$ be a topological space such that there exists one continuous function f on $X\times X$ to $[0, 1]$ such that $\mathfrak{I}$ is the least topology making f bi-equicontinuous. Then, $(X, \mathfrak{I})$ is called a minimal bi-equicontinuous space.*

REMARK. The metrizable topological spaces coincide with the minimal bi-equicontinuous spaces, by Corollary 1.1.

DEFINITION 3.2. *Let $(X, \mathfrak{I})$ be a topological space such that $\mathfrak{I}$ is the least topology making all continuous functions on $X\times X$ to $[0, 1]$ bi-equicontinuous. Then, $(X, \mathfrak{I})$ is called a maximal bi-equicontinuous space.*

DEFINITION 3.3. *Let* $(X, \mathcal{U})$ *and* $(Z, \mathcal{V})$ *be uniform spaces. Let* Y *be a topological space. For each* $x \in X$, *let* f_x *be a function on* Y *to* Z. *Let* $A \subset X$. *Then,* $\{f_x : x \in A\}$ *is uniformly extensively equicontinuous at* $y \in Y$ *if and only if for each* $\alpha \in \mathcal{V}$ *there exists a neighborhood of* y, $N(y)$, *and* $\beta \in \mathcal{U}$, *such that*

$$(1) \qquad f_{x'}(N(y)) \subset \alpha[f_x(y)] \qquad \textit{for each } x \in A,\ x' \in \beta[x].$$

LEMMA 3.1. *Let* $(X, \mathcal{U})$, $(Z, \mathcal{V})$ *be uniform spaces. Let* Y *be a topological space. Let* f *be a continuous function on* $X \times Y$ *to* Z. *Let* K *be a compact subset of* X. *Then,* $\{f_x : x \in K\}$ *is uniformly extensively equicontinuous at each* $y \in Y$.

Proof. Let $\alpha \in \mathcal{V}$, $y \in Y$. Choose $\alpha_0 \in \mathcal{V}$, α_0 symmetric and such that $\alpha_0^2 = \alpha_0 \cdot \alpha_0 \subset \alpha$. Since f is continuous on $X \times Y$, for each $k \in K$ there exist $\beta_k \in \mathcal{U}$ and $N_k(y)$, a neighborhood of y, such that

(i) $f(\beta_k[k] \times N_k(y)) \subset \alpha_0[f(k, y)]$.

Choose $\gamma_k \in \mathcal{U}$, γ_k open in $X \times X$ and such that $\gamma_k^2 = \gamma_k \circ \gamma_k \subset \beta_k$ for each $k \in K$. Now, K is compact and $K \subset \bigcup_{k \in K} \gamma_k[k]$. Therefore, there exists $k_1, k_2, \cdots, k_n \in K$ such that

(ii) $K \subset \bigcup_{i=1}^{n} \gamma_{k_i}[k_i]$.

(iii) Let $N(y) = \bigcap_{i=1}^{n} N_{k_i}(y)$. Let $\beta = \bigcap_{i=1}^{n} \gamma_{k_i}$. Let $k \in K$.

There exists $i \in \{1, 2, \cdots, n\}$ such that $k \in \gamma_{k_i}[k_i]$. Therefore, $f(\beta[k] \times N(y)) \subset f(\gamma_{k_i}[k] \times N(y))$, by (iii), $\subset f(\gamma_{k_i}^2[k_i] \times N(y)) \subset f(\beta_{k_i}[k_i] \times N(y)) \subset \alpha_0[f(k_i, y)]$, by (i) $\subset \alpha_0^2[f(k, y)]$ since α_0 is symmetric, $\subset \alpha[f(k, y)]$.

LEMMA 3.2. *Let* X *be a compact space,* f *a continuous function on* $X \times X$ *into* $(Y, \mathcal{V})$, *a uniform space. If* α *is open in* $Y \times Y$, *then* $\bigcap_{x \in X} f_x^{-1}(\alpha[f_x(y)])$ *is open in* X, $y \in Y$.

Proof. Let $z \in \bigcap_{x \in X} f_x^{-1}\alpha[f_x(y)]$. Then, $(f_x(y), f_x(z)) \in \alpha$, open, for each $x \in X$. Hence, there exists $\beta_x \in \mathcal{V}$ such that $\beta_x[f_x(y)] \times \beta_x[f_x(z)] \subset \alpha$ for each $x \in X$. Now, let $K = \{(f_x(y), f_x(z)) ; x \in X\}$. Then, K is compact since $K = (f^y, f^z)(\Delta_{X \times X})$ where $\Delta_{X \times X} \subset X \times X$ and $\Delta_{X \times X} = \{(x, y) : x = y\}$; also, $(f^y, f^z)(x, x) = (f^y(x), f^z(x)) = (f_x(y), f_x(z))$. $\Delta_{X \times X}$ is compact since X is compact and Hausdorff. By the uniform covering property of compact subspaces, there exists $\gamma \in \mathcal{V}$ such that for each $x \in X$, $\gamma[f_x(y)] \times \gamma[f_x(z)] \subset \alpha$. Thus,

(i) $\gamma[f_x(z)] \subset \alpha[f_x(y)]$ for each $x \in X$.

Now by Lemma 3.1, there exists a neighborhood of z, $N(z)$, such that $f_x(N(z)) \subset \gamma[f_x(z)]$, for each $x \in X$, $\subset \alpha[f_x(y)]$ for each $x \in X$, by (i). Thus, $N(z) \subset \bigcap_{x \in X} f_x^{-1}(\alpha[f_x(y)])$, and hence, $\bigcap_{x \in X} f_x^{-1}(\alpha[f_x(y)])$ is open.

COROLLARY 3.2. *Let* $(X, \mathfrak{I})$ *be a compact space,* f *a continuous function on* $X \times X$ *into* $(Y, \mathcal{V})$, *a uniform space. Let* $\mathfrak{I}_f$ *denote the topology on* X *which has as a base the sets,* $\{\bigcap_{x \in X} f_x^{-1}(\alpha[f_x(y)]) : \alpha$ *open,* $\alpha \in \mathcal{V}$, $y \in X\}$. *Then,* $\mathfrak{I}_f$ *is the least topology on* X *such that* $\{f_x : x \in X\}$ *is equicontinuous in the sense of Definition* 1.2. *Also,* $\mathfrak{I}_f \subset \mathfrak{I}$.

Proof. Let $\mathcal{A}$ be the open members of $\mathcal{V}$. By Lemma 3.2, for each $\alpha \in \mathcal{A}$,

$\bigcap_{x\in X} f_x^{-1}(\alpha[f_x(z)])$ is an open neighborhood of z in $\mathfrak{J}$. Hence, $\mathfrak{J}_f \subset \mathfrak{J}$. Also, $\{\bigcap_{x\in X} f_x^{-1}(\alpha[f_x(z)]) : \alpha \in \mathfrak{A}\}$ is a neighborhood base in $\mathfrak{J}_f$ of each $z \in X$. Let $\bigcap_{x\in X} f_x^{-1}(\alpha[f_x(y)])$ be a basic open set containing z. Then, by (i) in proof of Lemma 3.2, there exists $\gamma \in \mathfrak{V}$ such that $\bigcap_{x\in X} f_x^{-1}(\gamma[f_x(z)]) \subset \bigcap_{x\in X} f_x^{-1}(\alpha[f_x(y)])$.

REMARK. In Lemma 3.2 and Corollary 3.2, f_x may be replaced by f^x for each $x \in X$.

THEOREM 3.1. *Let* $(X, \mathfrak{J})$ *be compact. Then,* $\mathfrak{J}$ *is the least topology on* X *such that the set of all continuous functions on* $X \times X \to [0, 1]$ *is bi-equicontinuous; i.e.,* $(X, \mathfrak{J})$ *is a maximal bi-equicontinuous space.*

Proof. We use the fact that a compact space is uniformizable. By Corollary 2.1, there exist *some* continuous functions F on $X \times X$ to $[0, 1]$ such that $\mathfrak{J}$ is the least topology such that each $f \in F$ is bi-equicontinuous. If $\mathfrak{J}_G$ is the least topology on X such that for each $g \in G$, g continuous on $X \times X$ to $[0, 1]$, $\{g_x\}$ $[\{g^x\}]$ is equicontinuous, and if $F \subset G$, then $\mathfrak{J} \subset \mathfrak{J}_G$. Let A denote the set of *all* continuous functions on $X \times X$ to $[0, 1]$. $F \subset A$. Therefore, $\mathfrak{J} \subset \mathfrak{J}_A$. Also, by Corollary 3.2 and subsequent Remark, each $a \in A$ generates a topology $\mathfrak{J}_a \subset \mathfrak{J}$ which is the least topology such that $\{a_x\}$ $[\{a^x\}]$ is equicontinuous. Thus, $\mathfrak{J}_A \subset \mathfrak{J}$, and therefore $\mathfrak{J} = \mathfrak{J}_A$ or $\mathfrak{J}$ is the least topology such that the set of *all* continuous functions on $X \times X$ into $[0, 1]$ is bi-equicontinuous.

Reference

1. J. L. Kelley, General Topology, Van Nostrand, Princeton, N. J., 1955.

ARITHMETICAL FUNCTIONS AND DISTRIBUTIVITY

J. LAMBEK, McGill University, Montreal

By an *arithmetical* function will be understood a function from the positive integers to the complex numbers. Among various ways of combining such functions, two operations recommend themselves, ordinary multiplication and Dirichlet multiplication. We write

$$(fg)(n) = f(n)g(n),$$
$$(f \circ g)(n) = \sum_{n=dd'} f(d)g(d').$$

Both operations are commutative and associative. The neutral elements under these operations are the functions 1 and δ, where

$$1(n) = 1,$$
$$\delta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Thus $f1 = f$, $f \circ \delta = f$.

An arithmetical function f is called (completely) *multiplicative* if

$$f(dd') = f(d)f(d')$$

for all positive integers d and d'; (f is called *factorable* if this is true when d and d' are relatively prime). It is easily verified that every multiplicative function f is "distributive" in the sense that

$$f(g \circ h) = fg \circ fh$$

for all arithmetical functions g and h, and we shall leave the verification to the reader.

For example, let

$$\mu(n) = \begin{cases} (-1)^{k(n)} & \text{if } n \text{ is square-free,} \\ 0 & \text{otherwise,} \end{cases}$$

where $k(n)$ is the number of (distinct) prime factors of n. This is known as the Möbius function. Then $1 \circ \mu = \delta$, as is shown in almost any book on the theory of numbers. For multiplicative f it follows that

$$f \circ f\mu = f1 \circ f\mu = f(1 \circ \mu) = f\delta.$$

When moreover $f(1) = 1$, this becomes $f \circ f\mu = \delta$, and so $f\mu$ is the inverse of f under the operation $\circ$. The following appeared in [2].

THEOREM 1. *The arithmetical function f satisfies $f(g \circ h) = fg \circ fh$, for all arithmetical functions g and h, if and only if f is multiplicative.*

Proof. Assume that f is distributive.

CASE 1. $f(1) = 0$. For $k > 1$, let

$$\delta_k(n) = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{otherwise.} \end{cases}$$

Then $f(\delta_k \circ 1) = f\delta_k \circ f$, that is $f(n) \sum_{d|n} \delta_k(d) = \sum_{n=dd'} f(d)\delta_k(d)f(d')$. When $n = k$, this becomes $f(k)\delta_k(k) = f(k)\delta_k(k)f(1)$, that is, $f(k) = 0$. Thus $f(n) = 0$ for all n.

CASE 2. $f(1) \neq 0$. Let $\tau = 1 \circ 1$, so that $\tau(n)$ is the number of divisors of n. Then $f\tau = f(1 \circ 1) = f1 \circ f1 = f \circ f$, that is

$$f(n)\tau(n) = \sum_{n=dd'} f(d)f(d').$$

Taking $n = 1$, we obtain $f(1) = 1$. Taking $n = p_1 p_2 \cdots p_k$, the product of distinct primes, we shall prove by induction on k that

$$f(n) = f(p_1)f(p_2) \cdots f(p_k).$$

Indeed, the statement holds when $k = 1$. We assume the result for all $k' < k$,

where $k>1$. For $n=p_1p_2 \cdots p_k$ we have

$$2^k f(n) = f(n)\tau(n) = \sum_{n=dd'} f(d)f(d'),$$

hence $(2^k-2)f(n)=\sum' f(d)f(d')$, where $\sum'$ extends over all factorizations $n=dd'$ such that $d\neq n$ and $d'\neq n$. By the inductive hypothesis, this is

$$\sum_{1\leq r<k} \binom{n}{r} f(p_1) \cdots f(p_k) = (2^k-2)f(p_1) \cdots f(p_k),$$

hence $f(n)=f(p_1) \cdots f(p_k)$, as required.

Let us now consider any n, not necessarily a product of distinct primes, say $n=p_1^{r_1} \cdots p_k^{r_k}$. Define

$$f^*(n) = f(p_1^{r_1}) \cdots f(p_k^{r_k}).$$

Then f and f^* have the same values for square-free arguments, so that $f\mu=f^*\mu$, μ being the Möbius function. Since f and f^* are distributive, we have

$$f \circ f\mu = \delta, \qquad f^* \circ f^*\mu = \delta,$$

hence,

$$f = \delta \circ f = f^* \circ f^*\mu \circ f = f^* \circ f\mu \circ f = f^* \circ \delta = f^*,$$

and our argument is complete.

In spite of the above result, $(f \circ g)(h \circ k) \neq fh \circ fk \circ gh \circ gk$ in general, even when f, g, h and k are multiplicative. The correct law is the following.

THEOREM 2. *If f, g, h and k are multiplicative functions, then*

$$fh \circ fk \circ gh \circ gk = (f \circ g)(h \circ k) \circ u,$$

where

$$u(n) = \begin{cases} f(\sqrt{n})g(\sqrt{n})h(\sqrt{n})k(\sqrt{n}) & \text{when } n \text{ is a square,} \\ 0 & \text{otherwise.} \end{cases}$$

This result appeared in [2] also; however, with u replaced by its inverse under the operation $\circ$ on the other side of the equation. A similar result was also found by Vaidyanathaswamy.

My original proof of Theorem 2 went like this: We observe that both sides of the equation are factorable functions, hence it suffices to prove the result for an argument which is a prime power, etc. I now think that a direct proof of the theorem is more instructive.

Proof. Evaluating the left hand side at n, we obtain

$$\sum_{n=abcd} f(a)h(a)f(b)k(b)g(c)h(c)g(d)k(d) = \sum_{n=abcd} f(ab)g(cd)h(ac)k(bd)$$

$$= \sum_{n=xy=zt} f(x)g(y)h(z)k(t)\pi(x, y, z, t),$$

where $\pi(x, y, z, t)$ is the number of ways of writing $x=ab$, $y=cd$, $z=bc$, $t=bd$.

Evaluating the right hand side at n, we obtain

$$\sum_{n=de^2} \sum_{d=qr} f(q)g(r) \sum_{d=st} h(s)k(t)f(e)g(e)h(e)k(e)$$

$$= \sum_{n=de^2} \sum_{d=qr=st} f(qe)g(re)h(se)k(te)$$

$$= \sum_{n=xy=zt} f(x)g(y)h(z)k(t)\tau(x, y, z, t),$$

where $\tau(x, y, z, t)$ is the number of common divisors of x, y, z and t.

To complete the proof, we need only show that $\pi(x, y, z, t)=\tau(x, y, z, t)$. This result is of some small interest in itself, and we display it as a separate lemma.

As is well known, the diophantine equation $xy=zt$ has the parametric solutions $x=ab$, $y=cd$, $z=ac$, $t=bd$.

LEMMA. *Let $xy=zt$. Then there is a one-one correspondence between all quadruples (a, b, c, d) such that $x=ab$, $y=cd$, $z=bc$, $t=bd$, and all common divisors e of x, y, z, t.*

Proof. Given a common divisor e of x, y, z, t, we let

$$a = gcd(x/e, z/e),\ b = x/a,\ c = z/a,\ d = at/x.$$

(We omit the routine verification that $bd=t$, $cd=y$ and that d is an integer.) Then $gcd(b, c)=gcd(x/a, z/a)=e\cdot gcd(x/ae, z/ae)=e$. Conversely, if $e=gcd(b, c)$, then

$$gcd(x/e, z/e) = gcd(ab/e, ac/e) = a\cdot gcd(b/e, c/e) = a.$$

It is customary to associate with each arithmetical function f its *generating* function $\sum_{n\geq 1} f(n)n^{-s}$. For our purposes, it suffices to regard this as a formal infinite series, and we shall forget all about convergence. Then

$$\sum (f\circ g)(n)n^{-s} = \sum f(n)n^{-s} \sum g(n)n^{-s},$$

and Theorem 2 may be written

$$(\dagger)\qquad \begin{aligned} &\sum f(n)h(n)n^{-s} \sum f(n)k(n)n^{-s} \sum g(n)h(n)n^{-s} \sum g(n)k(n)n^{-s} \\ &\quad = \sum (f\circ g)(n)(h\circ k)(n)n^{-s} \sum f(n)g(n)h(n)k(n)n^{-2s}. \end{aligned}$$

Putting $f(n)=n^a$, $h(n)=n^b$, $g(n)=1=k(n)$, one gets Ramanujan's identity

$$\zeta(s-a-b)\zeta(s-a)\zeta(s-b)\zeta(s) = \zeta(2s-a-b) \sum \sigma_a(n)\sigma_b(n)n^{-s},$$

where $\zeta(s)$ is the Riemann zeta-function and $\sigma_a(n)$ is the sum of the ath powers of all divisors of n.

Dividing (†) by the formula obtained from (†) upon interchanging g and h, one gets

$$\frac{\sum (f \circ g)(n)(h \circ k)(n)n^{-s}}{\sum (f \circ h)(n)(g \circ k)(n)n^{-s}} = \frac{\sum f(n)h(n)n^{-s} \sum g(n)k(n)n^{-s}}{\sum f(n)g(n)n^{-s} \sum h(n)k(n)n^{-s}}$$

for multiplicative f, g, h and k, which may be of some interest also.

The author's interest in this problem was stimulated by Professor Gordon Pall.

References

1. E. T. Bell, Algebraic arithmetic, Amer. Math. Soc. Colloq. Publ., 7 (1927).

2. J. Lambek, A nondistributive calculus of numerical functions, McGill University, M.Sci. Thesis, 1946.

3. R. Vaidyanathaswamy, Multiplicative arithmetic functions, Trans. Amer. Math. Soc., 33 (1931) 579–662.

4. S. Ramanujan, Some formulae in the analytic theory of numbers, Messenger of Math., 45 (1916) 81–84. (Reprinted in "Collected Papers of S. Ramanujan," Chelsea, New York, 1962, pp. 133–135.)

MATHEMATICAL NOTES

Edited by J. H. Curtiss, University of Miami

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457

A REGULAR HOMOTOPY OF S^2

Richard Jerrard, University of Illinois, Urbana

Smale [1] has shown that any two immersions of S^2 in E^3 are regularly homotopic. It is interesting (and nontrivial) to construct a regular homotopy of S^2 in E^3 which carries S^2 into its mirror image, i.e., reverses the orientation. Such an orientation-reversing regular homotopy of S^1 in E^2 does not exist.

To construct such a regular homotopy, we must find a smoothly varying family of differentiable immersions f_t: $S^2 \to E^3$ $(0 \leqq t \leqq 1)$ which turns S^2 inside out. We start with a topological 2-sphere $S^2(0)$ which intersects a fixed cylinder C in two circles $A(0)$ and $B(0)$, as in Fig. 1a. The part of S^2 exterior to C is a half torus $T(0)$ bounded by $A(0)$ and $B(0)$ on C. The 2-sphere is completed by the two flat discs $D_A(0)$ and $D_B(0)$. We will put

$$f_t(S^2) = S^2(t) = T(t) + A(t) + B(t) + D_A(t) + D_B(t),$$

that is, for each value of t the sphere $S^2(t)$ consists of two open discs, two circles, and an open half torus.

The regular homotopy of S^2 will be defined by first constructing a regular homotopy of the circles A and B so that $A(1)=B(0)$, $B(1)=A(0)$, and $A(t)$, $B(t)$ lie on C for all t. This family of mappings will then be extended to all of S^2.

The immersions $A(t)$ and $B(t)$ of the circles on C will not have more than four self-intersections. It is shown below how to construct a smooth family of discs $D_A(t)$ such that $\partial D_A(t)=A(t)$ and $D_A(t)$ joins smoothly with $T(t)$. The half torus $T(0)$ can be generated by rotating a semi-circular arc around the cylinder C. For each t, the half torus $T(t)$ will be generated by moving a circular arc of smoothly varying radius around C with its endpoints on $A(t)$ and $B(t)$.

The illustrations in Figs. 1 and 2 (with the exception of Fig. 1a) show successive positions of $A(t)$ and $B(t)$ on C. The illustrated curves are supposed to be wrapped around C, with left and right endpoints identified. The sharp corners shown are for convenience of illustration; they are to be smoothly rounded in the actual mappings.

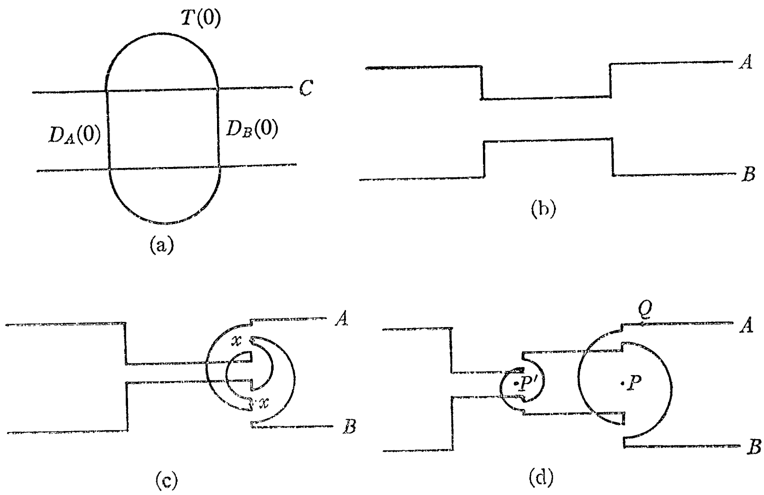


FIG. 1

The first step (Fig. 1b) is to pinch A and B together over about half of their length. The second step is to begin the reversal of orientation by the rotation shown in Fig. 1c. Throughout this process there corresponds to each point on $A(t)$ a symmetrically located point on $B(t)$. This is so because the curved lines in the figures are symmetric pairs with respect to centers of rotation, while the straight lines are symmetric pairs with respect to the center line between A and B.

In moving from Fig. 1c to 1d we simply extend the portions of A and B which have their positions interchanged, starting this elongation at the points

x, x in Fig. 1c. We then continue this process, leaving fixed the rotation centered at P while moving the rotation centered at P' around the cylinder, and at the same time expanding the size of this second rotation, arriving at the stage shown in Fig. 2a. The points P and Q, shown in Figs. 1d and 2a, remain fixed on C during this process, while the point P' has moved around the cylinder.

From Fig. 2a to Fig. 2b the rotation centered at P' is pushed further, until P and P' coincide. We arrive at a situation essentially the same as in Fig. 1c, except that the two circles A and B have almost finished interchanging their positions. All that remains is to untwist the complicated part of Fig. 2b, obtaining Fig. 2c, and then to unpinch the circles B and A. This process clearly defines a regular homotopy on C of the two circles A and B, taking $A(0)$ into $B(0)$ and $B(0)$ into $A(0)$.

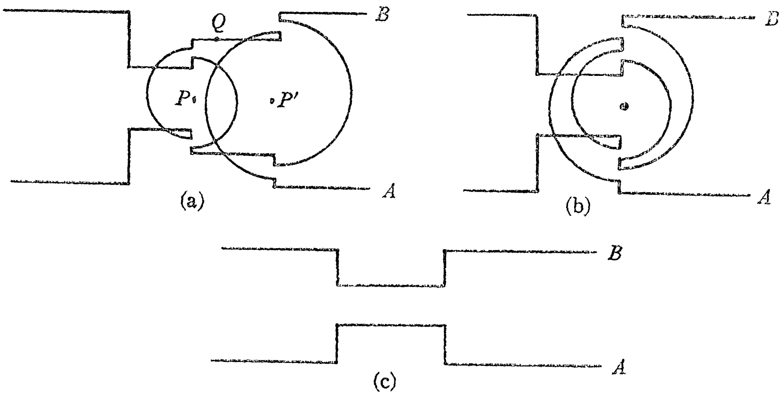


FIG. 2

The surface $T(t)$ is to consist of all points lying on circular arcs exterior to C whose endpoints are symmetric pairs on $A(t)$ and $B(t)$, and which intersect C orthogonally. There is just one such arc for each pair of points. The regular homotopy of A and B has been chosen so that the set of circular arcs comprising $T(t)$ defines a surface without cusps or other singularities, though of course it will have self-intersections. Thus a regular homotopy taking $T(0)$ into itself (with reversed orientation) is defined.

The appearance of the complicated part of $T(t)$ at a stage half way between Figs. 1b and 1c is shown in Fig. 3. The figure is not quite accurate because it shows the circular arcs generating $T(t)$ as semi-circles. This is almost the case, since we can assume that the radius of C is large compared to the radius of these arcs, i.e., locally the surface of C looks like a plane. The doubly curved portions of the surface in Fig. 3, taken separately, are half Möbius bands.

To show that the regular homotopy can be extended to the discs D_A and D_B, first parameterize with t the segment of the center line of C lying between the center of $A(0)$ and the center of $B(0)$. Denote these points by $K_A(t)$, with

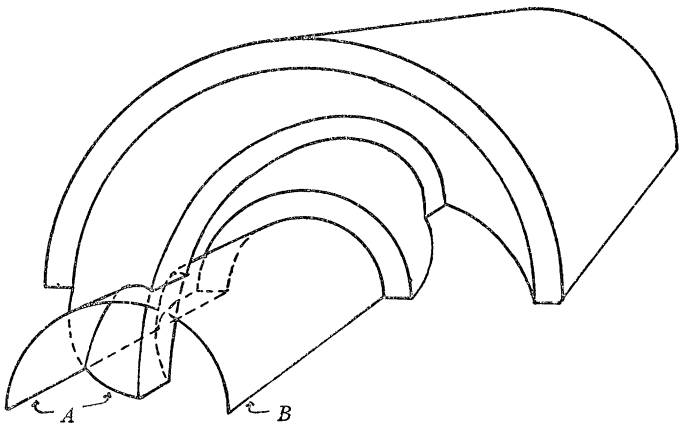


FIG. 3

$K_A(0)$, $K_A(1)$ being respectively the centers of $A(0)$ and $B(0)$. To generate $D_A(t)$ we can use a generating curve of suitable shape such as the sine curve of arbitrary amplitude between $-\pi$ and π. Then $D_A(t)$ consists of all such sine curves with endpoints on $A(t)$ and $K_A(t)$ which intersect C orthogonally and are tangent to a radius of C at $K_A(t)$. Each curve lies in a plane containing the center line of C; there is just one such curve for each point of $A(t)$. The smooth family of sine curves generates the required disc $D_A(t)$.

Finally we can suppose that the half torus $T(t)$ is generated by a circular arc moving in such a way that no cusps or other singularities appear. The arcs and curves generating the half torus and the discs all intersect C orthogonally. Therefore $S^2(t)$ intersects C orthogonally and is smooth at this intersection. We thus have a smooth family of differentiable immersions which accomplishes the desired results.

Reference

1. S. Smale, A classification of the immersions of the two-sphere, Trans. Am. Math. Soc., 90 (1959) 281–290.

EXTENSIONS OF THE SCHNIRELMANN DENSITY TO HIGHER DIMENSIONS

B. KVARDA AND R. KILLGROVE, San Diego State College

A definition of density for a set of positive integers has been given by Schnirelmann in [9] as follows. Let A be a set of positive integers, and for a given positive integer n, let $A(n)$ be the number of integers a in A such that $a \leqq n$. Then the *density* of A is the quantity

$$\alpha = \underset{n}{\text{g.l.b.}} \frac{A(n)}{n} \cdot$$

Clearly $0 \leqq \alpha \leqq 1$, and $\alpha = 1$ if and only if A is the set of all positive integers. It is also the case, as has been shown by Cheo in [2] and Lepson in [6], that for any real number α such that $0 \leqq \alpha \leqq 1$ there exists a set A of positive integers for which the density is α. We will restrict our attention to extensions of the Schnirelmann density which possess analogous properties.

For any sets A and B of positive integers, let C be the set of all integers of the form a, b, and $a+b$, where a is in A and b is in B, and let α, β, γ be the densities of A, B, C, respectively. Landau [5] and Schnirelmann [9] have proved

(1) If $\alpha + \beta \geqq 1$ then $\gamma = 1$, and (2) $\gamma \geqq \alpha + \beta - \alpha\beta$.

Mann [7] has strengthened (2) by proving the well-known $\alpha+\beta$ theorem,

$$\gamma \geqq \min(1, \alpha + \beta). \tag{3}$$

In [8] Mann has proved the following theorem, and has shown in a corollary that it implies (3):

Let n be any positive integer. Then either $C(n) = n$ or there exist positive integers m and m_1 not in C such that $m \leqq n$, $m_1 \leqq \max(m, n-m-1)$, and

$$\frac{C(n)+1}{n+1} \geqq \frac{A(m)+B(m)+1}{m+1} + \left| \frac{C(n)+1}{n+1} - \frac{C(m_1)+1}{m_1+1} \right| . \tag{4}$$

Now let Q_n be the set of all n-dimensional lattice points $(x_1, \cdots, x_n)$ for which each x_i is a nonnegative integer and at least one x_i is positive. If A is any subset of Q_n and T is any finite subset of Q_n, let $A(T)$ be the number of elements in $A \cap T$.

The set Q_2 may be identified with the set of nonzero Gaussian integers $x+yi$ for which x and y are both nonnegative rational integers. Cheo [1] has defined a density for subsets of Q_2 as follows, using our notation. Let x_0+y_0i be any element of Q_2, and let S be the set of all $x+yi$ in Q_2 such that $x \leqq x_0$, $y \leqq y_0$. Then for any subset A of Q_2, the density of A is the g.l.b.$_S[A(S)/Q_2(S)]$. Cheo has shown that (1) and, at least, a weaker version of (2) hold for his density. He has also given an example to show that (3) need not be true.

In [3] a *fundamental subset* of Q_n or, briefly, a *fundamental set*, is defined to be any finite nonempty subset R of Q_n such that whenever an element $(r_1, \cdots, r_n)$ is in R, we have also in R all elements $(x_1, \cdots, x_n)$ of Q_n with $x_i \leqq r_i$, $i = 1, \cdots, n$. Then for any subset A of Q_n, the density of A is defined in [3] to be the g.l.b.$_R[A(R)/Q_n(R)]$. For this density, (1) and (2) hold. Whether or not (3) is true is an open question, but in [4] a theorem analogous to Theorem II of [7] is proved.

We will now introduce a definition of density for subsets of Q_n for which a partial analogue to (4), hence also (1), (2), and (3), can easily be shown to be true. Accordingly, let $\bar{t} = (t_1, \cdots, t_n)$ be any lattice point in Q_n for which the components $t_1, \cdots, t_n$ have greatest common divisor 1, and let T be the set of all lattice points $k\bar{t} = (kt_1, \cdots, kt_n)$, where k is a positive integer. Note that if $n = 2$ or 3, T is just the set of all first quadrant or first octant lattice points on a

ray emanating from the origin. The set T will be called the *ray set* generated by $\bar{t}$.

Now let A be any subset of Q_n, and let $A(k\bar{t})$ be the number of lattice points $k'\bar{t}$ in $A\cap T$ for which $k'\leqq k$. We will define the density of A on T to be the quantity

$$\alpha_T = \operatorname*{g.l.b.}_{k} \frac{A(k\bar{t})}{k},$$

and the *ray density* of A to be the g.l.b.$_T$ α_T.

For any subsets A and B of Q_n, let C be the set of all lattice points $\bar{a}$, $\bar{b}$, $\bar{a}+\bar{b}$, where $\bar{a}$ is in A and $\bar{b}$ is in B, addition of lattice points being done componentwise. Then for any ray set T we see that $C\cap T$ contains all the elements in $A\cap T$ and in $B\cap T$, and all elements $\bar{a}+\bar{b}$ where $\bar{a}$ is in $A\cap T$ and $\bar{b}$ is in $B\cap T$, plus, possibly, additional points. Hence, because of the corresponding results for sets of positive integers, if $n\bar{t}$ is any element of T then either $C(n\bar{t})=n$ or there exist $m\bar{t}$, $m_1\bar{t}$ in T but possibly not in C such that $m\leqq n$, $m_1\leqq \max(m, n-m-1)$, and

$$\frac{C(n\bar{t})+1}{n+1} \geqq \frac{A(m\bar{t})+B(m\bar{t})+1}{m+1} + \left|\frac{C(n\bar{t})+1}{n+1} - \frac{C(m_1\bar{t}+1}{m_{1+1}}\right|.$$

Also, if α_T, β_T, γ_T are the densities on T of A, B, C, respectively, and if α, β, γ are the ray densities for A, B, C, then for all ray sets T, $\gamma_T\geqq \min(1, \alpha_T+\beta_T) \geqq \min(1, \alpha+\beta)$, and it follows that $\gamma\geqq\min(1, \alpha+\beta)$.

We conclude this note by comparing the sizes of the various densities defined for subsets of Q_n. If A is a subset of Q_2, then clearly the Cheo density of A is greater than or equal to the density defined in [3]. Now let A be any subset of Q_n, let α_1 be the density of A as defined in [3], and let α_2 be the ray density of A. We will prove

$$\alpha_1 \geqq \alpha_2. \tag{5}$$

We may establish (5) by showing that $A(R)/Q_n(R)\geqq\alpha_2$ for every fundamental set R. Accordingly, let R be an arbitrary fundamental set. Since R is finite, there are finitely many ray sets $T_1, T_2, \cdots, T_m$ which have nonempty intersections with R. Supposing T_i to be generated by $\bar{t}_i$, we may let

$$T_i \cap R = \{\bar{t}_i, 2\bar{t}_i, \cdots, k_i\bar{t}_i\}, \qquad i = 1, \cdots, m.$$

Then for each $i=1, \cdots, m$,

$$\frac{A(k_it_i)}{k_i} \geqq \alpha_2, \quad \text{or} \quad A(k_i\bar{t}_i) \geqq \alpha_2 k_i.$$

Hence $\sum_{i=1}^m A(k_i\bar{t}_i)\geqq\alpha_2\sum_{i=1}^m k_i$. But $A(R)=\sum_{i=1}^m A(k_i\bar{t}_i)$, and $Q_n(R)=\sum_{i=1}^m k_i$. Consequently, $A(R)\geqq\alpha_2 Q_n(R)$ and $A(R)/Q_n(R)\geqq\alpha_2$, which completes the proof.

The authors would like to thank the referee for bringing to their attention Mann's stronger 1960 result.

References

1. L. P. Cheo, On the density of sets of Gaussian integers, this MONTHLY, 58 (1951) 618–620.
2. ———, A remark on the $\alpha+\beta$ theorem, Proc. Amer. Math. Soc., 3 (1952) 175–177.
3. B. Kvarda, On densities of sets of lattice points, Pacific J. Math., 13 (1963) 611–615.
4. ———, An inequality for the number of elements in a sum of two sets of lattice points, Pacific J. Math., 15 (1965) 545–550.
5. E. Landau, Über einige neuere Fortschritte der additive Zahlentheorie, Cambridge Univ. Press, London, 1937.
6. B. Lepson, Certain best possible results in the theory of Schnirelmann density, Proc. Amer. Math. Soc., 1 (1950) 592–594.
7. H. B. Mann, A proof of the fundamental theorem on the density of sums of sets of positive integers, Ann. Math., 43 (1942) 523–527.
8. ———, A refinement of the fundamental theorem on the density of the sum of two sets of integers, Pacific J. Math., 10 (1960) 909–915.
9. L. Schnirelmann, Über additive Eigenschaften von Zahlen, Math. Ann., 107 (1933) 649–690.

THE NUMBER OF TERMS IN THE CYCLOTOMIC POLYNOMIAL $F_{pq}(x)$

L. CARLITZ, Duke University

Let p, q denote distinct primes. Then the cyclotomic polynomial $F_{pq}(x)$ is given by

$$F_{pq}(x) = \frac{(1 - x^{pq})(1 - x)}{(1 - x^p)(1 - x^q)} \cdot \tag{1}$$

Bang [1] and Migotti [4] have proved that the coefficients of $F_{pq}(x)$ are ± 1 or 0. In a recent paper [2] Sister Marion Beiter has proved that if $F_{pq}(x) = \sum_{n=0}^{\phi(pq)} c_n x^n$, then

$$c_n = \begin{cases} (-1)^\delta & \text{if } n = \alpha q + \beta p + \delta \text{ in exactly one way,} \\ 0 & \text{otherwise,} \end{cases}$$

where α, β are nonnegative integers and $\delta = 0, 1$. She also determined the middle coefficient c_n, where $n = \phi(pq)/2$.

It is of some interest to determine the number of nonzero terms in $F_{pq}(x)$. Let $\theta_0(pq)$ denote the number of terms with positive coefficients, $\theta_1(pq)$ the number of terms with negative coefficients and $\theta(pq) = \theta_0(pq) + \theta_1(pq)$, the total number of nonzero terms. Since $F_{pq}(1) = 1$ it follows at once that $\theta_0(pq) = 1 + \theta_1(pq)$, so that

$$\theta(pq) = 2\theta_0(pq) - 1. \tag{2}$$

We may assume $q > p$ and define u by means of

$$qu \equiv -1 \pmod{p} \qquad (0 < u < p). \tag{3}$$

Then by (1) we have

$$F_{pq}(x) = \frac{1-x}{1-x^p}\sum_{j=0}^{p-1} x^{jq} = \frac{1}{1-x^p}\left\{\sum_{j=0}^{p-1} x^{jq} - \sum_{i=0}^{p-1} x^{iq+1}\right\}.$$

With each term x^{jq} of the first sum on the extreme right, associate the term x^{iq+1} of the second sum for which $iq+1 \equiv jq \pmod{p}$. By (3) it follows that $i \equiv j+u \pmod{p}$. Then

$$\begin{aligned} F_{pq}(x) &= \frac{1}{1-x^p}\sum_{j=0,j+u<p}^{p-1}(x^{jq} - x^{(j+u)q+1}) + \frac{1}{1-x^p}\sum_{j=0,j+u\geqq p}^{p-1}(x^{jq} - x^{(j+u-p)q+1}) \\ &= \frac{1-x^{uq+1}}{1-x^p}\sum_{j=0}^{p-1-u} x^{jq} - \frac{1-x^{(p-u)q-1}}{1-x^p}\sum_{i=0}^{u-1} x^{iq+1}. \end{aligned} \tag{4}$$

The first sum in (4) consists of the terms of $F_{pq}(x)$ with positive coefficients; there are evidently $(p-u)(uq+1)/p$ such terms. The second sum in (4) accounts for the negative terms of $F_{pq}(x)$; there are $u(pq-uq-1)/p$ such terms. Note that

$$\frac{(p-u)(uq+1)}{p} - \frac{u(pq-uq-1)}{p} = 1,$$

as we should expect.

This proves the following

THEOREM. *Let $\theta_0(pq)$ denote the number of terms with positive coefficients in $F_{pq}(x)$. Take $q>p$ and define u by means of $qu \equiv -1 \pmod{p}$, $(0<u<p)$. Then we have*

$$\theta_0(pq) = (p-u)(uq+1)/p. \tag{5}$$

For example if $p=3$ and $q=3k+1$, then $u=2$ and we get

$$\theta_0(3(3k+1)) = 2k+1.$$

If $p=3$ and $q=3k+2$, then $u=1$, so that $\theta_0(3(3k+2))=2k+2$. It is not difficult to verify these results directly. Similarly for $p=5$ we have

$$\theta_0(5(5k+1)) = 4k+1, \qquad \theta_0(5(5k+3)) = 6k+4,$$
$$\theta_0(5(5k+2)) = 6k+3, \qquad \theta_0(5(5k+4)) = 4k+4.$$

We remark that if we assume only that p, q are relatively prime but otherwise arbitrary integers greater than one, the above theorem continues to hold. The notation is somewhat ambiguous; to avoid confusion we may write $F_{p,q}(x)$ in place of $F_{pq}(x)$ and $\theta_0(p, q)$ in place of $\theta_0(pq)$. Then we have, for example,

$$\theta_0(4, 4k+1) = 3k+1,$$
$$\theta_0(4, 4k+3) = 3k+3.$$

Also for arbitrary p we have

$$\theta_0(p, kp+1) = k(p-1)+1,$$

$$\theta_0(p,\, kp + p - 1) = k(p - 1) + p - 1.$$

For p odd, on the other hand, we get

$$\theta_0(p,\, kp + 2) = \tfrac{1}{4}k(p^2 - 1) + \tfrac{1}{2}(p + 1),$$
$$\theta_0(p,\, kp + p - 2) = \tfrac{1}{4}(k + 1)(p^2 - 1) - \tfrac{1}{2}(p - 1).$$

The last four formulas indicate how strongly the value of $\theta_0(p, q)$ depends on the residue of q (mod p).

The following remark was suggested by the referee. A recent problem in this MONTHLY (E1637, 1964, p. 799) asks that, given $(a, b)=1$, $a>1$, $b>1$, one finds $N(a, b)$ the smallest positive integer such that for every $n \geqq N(a, b)$, there exists $x, y \geqq 0$ such that n is *representable* as $n=ax+by$. Similar questions may be found in [3], p. 22, problem 4 and [5], p. 97, problem 4.

LeVeque observes that the number of nonnegative integers less than $N(a, b)=(2-1)(b-1)$ which are *representable* in this manner is $(a-1)(b-1)/2$. One may ask how they are clustered; specifically, into how many sets of *consecutive* integers are the *representable* integers divided. It is readily seen that $c_n=+1$ if and only if n is the smallest of a set of consecutive representable integers while $c_n=-1$ if and only if n is the smallest of a set of consecutive *nonrepresentable* integers. Some sets contain just one element.

Supported in part by NSF grant GP-1593.

References

1. A. S. Bang, Om Ligningen $\phi_n(x)=0$, Nyt Tidsskrift for Mathematik, 6 (1895) 6–12.

2. Sister Marion Beiter, The midterm coefficient of the cyclotomic polynomial $F_{pq}(x)$, this MONTHLY, 71 (1964) 769–770.

3. W. J. LeVeque, Topics in Number Theory, vol. 1, Addison-Wesley, Reading, Mass., 1956.

4. A. Migotti, Zur Theorie der Kreisteilungsgleichung, S.-B. der Math.-Naturwiss. Classe der Kaiserlichen Akademie der Wissenschaften, Wien, (2) 87 (1883) 7–14.

5. I. Niven and H. S. Zuckerman, An Introduction to the Theory of Numbers, Wiley, New York, 1960.

ON GROUPS AND THEIR NEAR-RINGS OF FUNCTIONS

J. C. BEIDLEMAN, University of Kentucky

Introduction. A near-ring R is an algebraic system with two binary compositions of addition and multiplication such that:

1. R is a group under addition.
2. R is a semi-group under multiplication.
3. $a(b+c)=ab+ac$ for all a, b, and c from R.

Let G be an additive group (not necessarily abelian) and $R(G)$ the set of all functions on G. If the elements of $R(G)$ are added by adding images and multiplication is iteration, then $R(G)$ is a near-ring. Let $R_0(G)$ denote the set of all

functions on G that commute with the zero function. Then $R_0(G)$ is a sub-near-ring of $R(G)$.

Let $R'(G)=\{\phi_g\in R(G)\mid g'\phi_g=g;\ g,\ g'\in G\}$. Then $R'(G)$ is a sub-near-ring of $R(G)$ and the mapping $\eta g: \in G\to\phi_g\in R'(G)$ is a group isomorphism of the additive group G onto the additive group of $R'(G)$. If r is an element of $R(G)$, then it is easily verified that $r\phi_g=\phi_g$ and $\phi_g r=\phi_{gr}$. We now turn our attention to a sub-near-ring of $R_0(G)$. Define an element Ψ_g in $R_0(G)$ as follows:

$$g'\Psi_g = \begin{Bmatrix} g \text{ if } g' \neq 0 \\ 0 \text{ if } g' = 0 \end{Bmatrix} \quad \text{where } g,\ g' \in G.$$

If $R_0'(G)=\{\Psi_g\in R_0(G)\mid g\in G\}$, then $R_0'(G)$ is a sub-near-ring of $R_0(G)$ and its additive group is isomorphic to G. Throughout this article the near-rings $R_0(G)$ and $R_0'(G)$ will play an important role.

The main result of the first section is that two groups G and H are isomorphic if, and only if, there exists a near-ring isomorphism of $R(G)$ onto $R(H)$. In the second section we introduce the concept of kernel-preserving homomorphism and prove that G and H are isomorphic if, and only if, there exists a kernel-preserving near-ring isomorphism of $R_0(G)$ onto $R_0(H)$.

1. An isomorphism theorem.

(1.1) LEMMA. *Let f be a group isomorphism of G onto H. Then f induces a near-ring isomorphism of $R(G)$ onto $R(H)$.*

Proof. If r is an element of $R(G)$, then $f^{-1}rf$ is contained in $R(H)$. A simple calculation shows that the mapping $\eta: r\in R(G)\to f^{-1}rf\in R(H)$ is a near-ring isomorphism of $R(G)$ onto $R(H)$.

(1.2) LEMMA. *If $R(H)$ is a (near-ring) homomorphic image of $R(G)$, then H is a homomorphic image of G.*

Proof. Let f be a near-ring homomorphism of $R(G)$ onto $R(H)$. It suffices to prove $R'(G)f=R'(H)$. Let ϕ_g be an element of $R'(G)$ and ϕ_h an element of $R'(H)$. Then there exists an element r contained in $R(G)$ such that $rf=\phi_h$. Therefore, we obtain the following:

$$(1) \qquad (\phi_g)f = (r\phi_g)f = (rf)(\phi_g f) = \phi_h\cdot(\phi_g f) \in R'(H)$$

and

$$(2) \qquad (\phi_{gr})f = (\phi_g r)f = (\phi_g f)\cdot(rf) = (\phi_g f)\cdot\phi_h = \phi_h.$$

From (1) and (2) the lemma follows.

By lemmas (1.1) and (1.2) we have:

(1.3) THEOREM. *Two groups G and H are isomorphic if, and only if, there exists a near-ring isomorphism of $R(G)$ onto $R(H)$.*

2. Kernel preserving mappings. Let η be a function of G into H. By the kernel of η is meant $\mathrm{Ke}(\eta)=\{g\in G \mid g\eta=0\}$. We shall say that η is *kernel-free* if $\mathrm{Ke}(\eta)=0$. A near-ring homomorphism of $R_0(G)$ into $R_0(H)$ is called *kernel-preserving* if it maps kernel-free elements of $R_0(G)$ onto kernel-free elements of $R_0(H)$.

The nonzero elements of $R_0'(G)$ are kernel-free mappings of G into G. Let r be an element of $R_0(G)$. A routine calculation shows that r is kernel-free if, and only if, $r\Psi_g=\Psi_g$ for every nonzero element $g\in G$.

(2.1) THEOREM. *Two groups G and H are isomorphic if, and only if, there exists a near-ring isomorphism of $R_0(G)$ onto $R_0(H)$ which is kernel-preserving.*

Proof: Let f be a group isomorphism of G onto H. Then the mapping $\eta: r\in R_0(G)\to f^{-1}rf\in R_0(H)$ is a near-ring isomorphism of $R_0(G)$ onto $R_0(H)$. Let r be a kernel-free element of $R_0(G)$ and let h be an element of H such that $h(f^{-1}rf)=0$. Then $h(f^{-1}r)\in \mathrm{Ke}(f)=0$ and so it follows that $(hf^{-1})r=0$. Since r is kernel-free $h\in\mathrm{Ke}(f^{-1})=0$. This shows that η is kernel-preserving.

Assume that η is a near-ring isomorphism of $R_0(G)$ onto $R_0(H)$ which is kernel-preserving. We shall show that η induces a group isomorphism of the additive group of $R_0'(G)$ onto the additive group of $R_0'(H)$. Let Ψ_h be a nonzero element of $R_0'(H)$. Then there exists a kernel-free element $r\in R_0(G)$ such that $r\eta=\Psi_h$. If Ψ_g is a nonzero element of $R_0'(G)$, then we obtain

$$\Psi_g\eta = (r\Psi_g)\eta = (r\eta)\cdot(\Psi_g\eta) = \Psi_h\cdot(\Psi_g\eta)\in R_0'(H) \tag{1}$$

and

$$(\Psi_{gr})\eta = (\Psi_g\eta)\cdot(r\eta) = (\Psi_g\eta)\cdot\Psi_h = \Psi_h\in R_0'(H). \tag{2}$$

From (1) and (2) we conclude that $[R_0'(G)]\eta=R_0'(H)$ and so the theorem follows.

A NOTE ON GEOMETRY IN A p-RING

R. A. MELTER, University of Massachusetts

In a note in this MONTHLY [2], J. L. Zemmer exhibited a Boolean geometry (for the integers) in which every motion was of the form $m(x)=x+m(0)$ or $m(x)=-x+m(0)$. We shall call the group of all such motions the *geometric* group of the integers. Subsequent work [1, 3] on the Boolean metric space of a p-ring developed that the groups of motions for these spaces were quite different. In the present note we introduce the *operator space* of a p-ring R and show that each of its motions can be uniquely represented as the product of an element of the geometric group of the integers and a motion of the Boolean metric space of R.

It is well known that every p-ring is a subdirect sum of $\mathrm{GF}(p)$. We limit our attention to a particular p-ring: the discrete direct sum, that is, the ring of functions from the set of integers to $\mathrm{GF}(p)$ which take on nonzero values for at most a finite number of integers. Henceforth R will designate only this ring. Let i be an integer, f an element of R, and g an element of the geometric group of the

integers. The mapping of R onto itself, $f \to g(f)$ defined by $\{g(f)\}(i) = f(g(i))$ is an automorphism of R; the collection of all such automorphisms forms a group G which is isomorphic to the geometric group of the integers. Let $[x]$ be the equivalence class, under G, which contains x.

We note that the automorphisms described above are also automorphisms of the Boolean ring of idempotents B. Thus if x and y are elements of R we can define the distance between x and y to be $[(x-y)^{p-1}]$, where the equivalence class is now a subset of B. With this *distance* R becomes an abstract semi-metric space which we call its *operator space.*

THEOREM. *μ is a motion of the operator space of R if and only if it has a unique representation in the form $\mu = g\gamma$ where $g \in G$ and $\gamma \in \Gamma$, the group of motions of the Boolean metric space of R.*

Proof. $g\gamma$ is a motion of the operator space since

$$\begin{aligned} d(g\gamma(x), g\gamma(y)) &= [(g\gamma(x) - g\gamma(y))^{p-1}] = [(g(\gamma(x) - \gamma(y)))^{p-1}] \\ &= [g((\gamma(x) - \gamma(y))^{p-1})] = [g((x - y)^{p-1})] \\ &= [(x - y)^{p-1}] = d(x, y). \end{aligned}$$

It remains to show that every motion of the operator space has a unique representation in the form $\mu = g\gamma$, where $g \in G$ and $\gamma \in \Gamma$. Suppose that $\mu(0) = 0$. It follows that $[(\mu(x))^{p-1}] = [x^{p-1}]$ and hence that $(\mu(x))^{p-1} = g(x^{p-1})$, where g depends upon x. In Lemma 1 we show that for a certain subset of R, g can be chosen independent of x. It will follow from Lemma 2 that g can be chosen independent of x for all of R.

Denote by x_α any element of R which has a nonzero entry in its α-th co-ordinate and zeros elsewhere. If, in particular this nonzero entry is one, the element of R will be designated e_α. Let P be the set of all elements of R of the form x_α, where α is allowed to range over the integers.

LEMMA 1. *Let μ be a motion of the operator space such that $\mu(0) = 0$. There is a unique $g \in G$ such that $(\mu(x))^{p-1} = g(x^{p-1})$ for all $x \in P$.*

Proof of Lemma 1. Since g is a one-to-one map on the integers it follows that $(\mu(x_\alpha))^{p-1} = e_\beta$ for some integer β which is independent of the particular entry which x_α has in its α-th co-ordinate. Suppose that $(\mu(x_0))^{p-1} = e_w$; g, considered as a mapping on the integers has the property $|g(m) - g(n)| = |m - n|$. Hence exactly one of the two following sets of equations must hold.

$$(1) \qquad (\mu(x_{-1}))^{p-1} = e_{w-1}; \qquad (\mu(x_1))^{p-1} = e_{w+1}$$

$$(2) \qquad (\mu(x_1))^{p-1} = e_{w+1}; \qquad (\mu(x_1))^{p-1} = e_{w-1}.$$

In both cases there is a unique g such that $(\mu(x_s))^{p-1} = g((x_s)^{p-1})$ for $s = -1, 0, 1$. In the first case we can use $g(n) = n - w$, while in the second case the necessary mapping is $g(n) = -n + w$. Suppose we have shown that $(\mu(x_s))^{p-1} = g(x_s)^{p-1}$

for $-m \leqq s \leqq m$, where $g(n) = n - w$. It follows that

$$(\mu(x_{m+1}))^{p-1} = e_{w+m+1} \quad \text{and} \quad (\mu(x_{-m-1}))^{p-1} = e_{w-m-1}.$$

Hence $(\mu(x_s))^{p-1} = g((x_s)^{p-1})$ for $-(m+1) \leqq s \leqq m+1$, and by induction for all integers s. A similar argument for $g(n) = -n + w$ completes the proof of Lemma 1.

LEMMA 2. *Let μ be a motion of the operator space such that $\mu(0) = 0$. There is a unique $g \in G$ such that $(\mu(x) - \mu(y))^{p-1} = g((x-y)^{p-1})$ for all x, y in R.*

Proof of Lemma 2. Let g be the element of G provided by Lemma 1. We shall show that it also satisfies the equation of Lemma 2. If it does so it will certainly be unique since Lemma 1 is a special case of Lemma 2.

Suppose that Lemma 2 is false. Then there exist elements x and y in R such that $(\mu(x) - \mu(y))^{p-1} \neq g((x-y)^{p-1})$. Let δ be one of the co-ordinates in which these two idempotents differ. We will consider the case in which $(\mu(x) - \mu(y))^{p-1}$ has a 0 in the δ-th co-ordinate and $g((x-y)^{p-1})$ has a 1 in the δ-th co-ordinate. (The alternative case can be handled similarly.) It follows that $\mu(x)$ and $\mu(y)$ have the same entry in the δ-th co-ordinate while x and y have different entries in the β-th co-ordinate, where $\beta = g^{-1}(\delta)$. Let $\bar{x} = xe_\beta$ and $\bar{y} = ye_\beta$. It is a consequence of Lemma 1 that $\mu(\bar{x})$ and $\mu(\bar{y})$ will have different entries in the δ-th co-ordinate. An examination of the following pairs of idempotents

$$\begin{array}{ll} (\mu(x) - \mu(\bar{x}))^{p-1}, & (x - \bar{x})^{p-1} \\ (\mu(y) - \mu(\bar{y}))^{p-1}, & (y - \bar{y})^{p-1} \\ (\mu(x) - \mu(\bar{y}))^{p-1}, & (x - \bar{y})^{p-1} \\ (\mu(y) - \mu(\bar{x}))^{p-1}, & (y - \bar{x})^{p-1} \end{array}$$

reveals, however, that for at least one of the pairs the two members of the pair contain different numbers of ones, which is impossible if μ is a motion. This contradiction completes the proof of Lemma 2.

It follows from Lemma 2 that if μ is a motion which fixes 0, there is an element $g \in G$ such that $\gamma = g^{-1}\mu$ is a motion of the operator space with the property $(\gamma(x) - \gamma(y))^{p-1} = (x-y)^{p-1}$, i.e., γ is a motion of R considered as a Boolean metric space. Thus $\mu = g\gamma$ where $g \in G$ and $\gamma \in \Gamma$.

If μ does not fix 0, let ν map x into $x + \mu^{-1}(0)$. It is easy to see that ν is an element of Γ. It follows that $\sigma = \mu\nu$ is a motion of the operator space which maps 0 into 0 and hence $\sigma = g'\gamma' = \mu\nu$. Thus $\mu = g'\gamma'\nu^{-1} = g'\theta$, where $\theta \in \Gamma$ and $g' \in G$.

The proof has already established that if μ is a motion which fixes 0 and $\mu = g\gamma = g'\gamma'$, then $g = g'$ and hence $\gamma = \gamma'$. If a motion ρ which does not fix zero had two distinct representations then the motion $\rho\nu$ (where ν takes x into $x + \rho^{-1}(0)$) would fix 0 *and* have two distinct representations contrary to what has already been established.

The author would like to thank the Institute for Advanced Study for allowing him to visit the Institute during the period in which this note was written. He would also like to thank the referee and Professor Wallace Martindale for their interest.

References

1. J. L. Zemmer, Some remarks on p-rings and their Boolean geometry, Pacific J. Math., 6 (1956) 193–208.

2. ———, A Boolean geometry for the integers, this MONTHLY, 67 (1960) 56–57.

3. R. A. Melter, Contributions to Boolean geometry of p-rings, Pacific J. Math., 14 (1964) 995–1017.

ON REPEATED INTERCHANGE GRAPHS

V. V. MENON, Indian Statistical Institute, Calcutta (Presently at Trinity College, Oxford)

1. In the following, we will understand by a graph G a collection V of vertices, and a collection $\mathcal{A}$ of unordered pairs of vertices called edges. The edges need not be distinct (i.e., multiple edges can exist) and loops (i.e., edges of the form (v, v)) are allowed. For an edge (v_i, v_j), the vertices v_i and v_j are called its end-points, and v_i is said to be joined by this edge to v_j.

The Interchange Graph $I(G)$ of a graph G is obtained as follows: the edges of G form the vertex set of $I(G)$, and two vertices of $I(G)$ are joined by zero, one, or two edges according as the two corresponding edges in G have zero, one, or two end-points in common. In the literature, $I(G)$ is known also as the adjoint or the line-graph of G. In Figure 1 are given some graphs G_i and their interchange graphs $I(G_i)$. The vertices are labelled as $v_1, v_2, \cdots$, and the edges as $e_1, e_2, \cdots$.

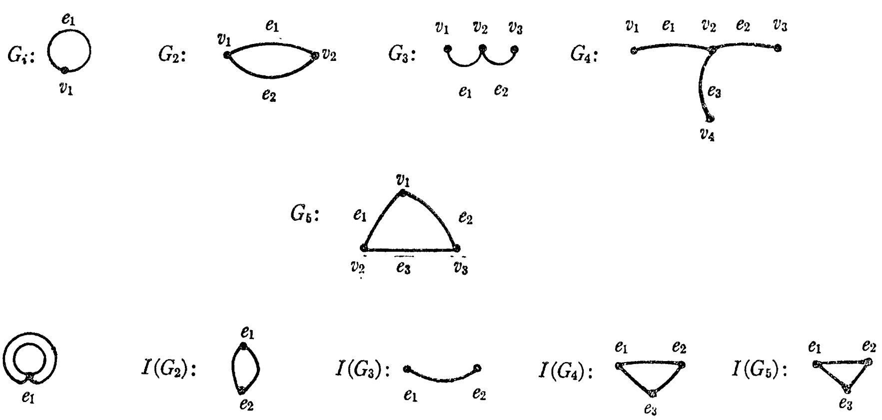


FIG. 1

Repeated interchange graphs are defined by the relation

$$I^n(G) = I[I^{n-1}(G)], \quad n = 2, 3, \cdots.$$

The object of our study is to examine the 'convergence' of the sequence $I^n(G)$, $n=1, 2, \cdots$. Ore [1] has posed a more general problem. Before proceeding, we introduce the following terminology.

A subgraph of a graph $G=[V; \mathcal{A}]$ is a graph $G_u=[U, \mathcal{B}]$ whose vertices form a subset of V, and whose edges form a subcollection of the edges in G whose end-

points are in U. A *connected component* of $G=[V, \mathfrak{A}]$ is a subgraph $G_u=[U, \mathfrak{B}]$ having the following properties:

(1) $\mathfrak{B}$ is the collection of all the edges in G having their end-points in U;

(2) any two vertices in U are joined by a chain of vertices and edges in G_u (i.e., if u_0 and u_n are the vertices, then there is a sequence of vertices $u_1, u_2, \cdots, u_{n-1}$ in U such that (u_i, u_{i+1}) is in $\mathfrak{B}$, for $i=0, 1, \cdots, n-1$);

(3) U is a maximal set having this property. The graph G is *connected* if G is the only connected component in it.

The *degree of a vertex* v is the number of edges one of whose end-points is v. The *degree of an edge* (v_i, v_j) is defined as d_i+d_j-2 where d_i is the degree of the vertex v_i, and d_j that of v_j. (In fact, in $I(G)$, this is the degree of the vertex corresponding to the edge (v_i, v_j) of G.)

The *0-graph* is a graph whose vertex set (and hence the edge set) is empty.

A graph $G=[V, \mathfrak{A}]$ is *finite* if the sum of the cardinal numbers of V and $\mathfrak{A}$ is finite.

We are dealing with finite graphs only.

For a finite graph G we denote by n_k and m_k the number of vertices and edges, respectively, of $I^k(G)$. If the sequence n_k+m_k, $k=1, 2, \cdots$ is bounded, we say that the sequence $I^k(G)$ *converges*; in all other cases, we say that $I^k(G)$, $k=1, 2, \cdots$ *diverges*.

We shall establish that the sequence $I^k(G)$ converges if and only if the degree of each vertex is at most two. In such a case, moreover, the sequence is stabilized, in the sense that $I^k(G)=I^n(G)$ for $k\geqq n_0$ where n_0 is the number of vertices in G. The equality of graphs $G=G'$, means that they are isomorphic, i.e., there is a one-one correspondence between the vertex sets of G and G' such that the number of edges joining any two vertices of G is equal to the number of edges joining the corresponding vertices in G'.

2. Let G_u be a subgraph (respectively, a connected component) of a finite graph $G=[V, \mathfrak{A}]$. Then $I(G_u)$ is a subgraph (respectively, a connected component) of $I(G)$. We can, therefore, assume that the graph G is connected (by considering the connected components separately). There are four cases to be considered.

Case 1. There are no loops, and the degree of each vertex is at most 2, there being at least one vertex whose degree is 0 or 1. In this case, we can write (since G is connected) the vertex set $V=\{v_1, v_2, \cdots, v_n\}$ and the edges as (v_i, v_{i+1}), $i=1, \cdots, n-1$. $I^{n-1}(G)$ consists of single vertex and no edges, and $I^n(G)$ is the 0-graph.

Case 2. There are no loops and the degree of each vertex is 2. Then we can write the vertex set as $V=\{v_1, v_2, \cdots, v_n\}$ and the edges as (v_n, v_1), (v_i, v_{i+1}), $i=1, 2, \cdots, n-1$. It is easily verified that $G=I(G)$ and hence $G=I^n(G)$ for all n.

Case 3. There are no loops and the degree of at least one vertex is at least 3. If there is a vertex v of degree $\geqq 4$, we can obtain a subgraph G_v of the form

shown in Figure 2. (Some of these four edges may join the same end-points.)

In $I(G_v)$, every vertex has a degree $\geqq 3$.

By a simple argument, one shows that if $d_1, d_2, \cdots, d_{n_0}$ are the degrees of the vertices in G, then there are $\frac{1}{2}\sum_{i=1}^{n} d_i(d_i-1)$ edges in $I(G)$ if G has no loops. (The edges at the vertex v_i of G contribute $\frac{1}{2}d_i(d_i-1)$ edges in $I(G)$.) Denote the number of vertices and edges in G by n_0 and m_0 respectively, and those in $I(G)$ by n_1 and m_1, respectively. Then

$$n_1 = m_0 = \frac{1}{2}\sum_{i=1}^{n_0} d_i, \qquad m_1 = \frac{1}{2}\sum_{i=1}^{n_0} d_i(d_i - 1).$$

It follows that if the degree of each vertex of G is $\geqq 3$, then the degree of each vertex in $I(G)$ also is $\geqq 3$, $n_1 > n_0$ and $m_1 > m_0$. By induction, therefore, $n_{k+1} > n_k$, and $m_{k+1} > m_k$.

Hence $I^k(G_v)$, $k = 1, 2, \cdots$, diverges.

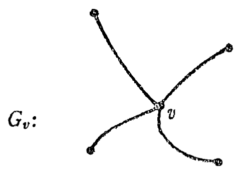


FIG. 2

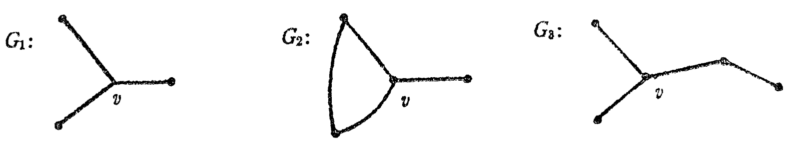


FIG. 3

On the other hand, if there is no vertex of degree $\geqq 4$, we take a vertex of degree 3. It is possible then to obtain at least one of the subgraphs shown in Figure 3.

One verifies that $I(G_1) = I^n(G_1)$ for all n, and $I^2(G_2) = I^3(G_3)$, each containing a vertex of degree 4. The same argument as in the above paragraph applies to $I^2(G_2)$ and $I^3(G_3)$. Of course, if the subgraphs G_2 or G_3 do not exist in G, then $G = G_1$.

Case 4. There is at least one loop in G. Consider the subgraph G_0 (shown as G_1 in Figure 1) consisting of a single loop. $I(G_0)$ consists of a vertex with two loops. $I^2(G_0)$ has two vertices joined by two edges and with two loops at each vertex, i.e., $I^2(G_0)$ has two subgraphs like $I(G_0)$. The number of vertices and of edges in $I^n(G)$ increases as n increases.

3. We notice that if a subgraph of G diverges, then G itself diverges. The material of Section 2 can then be stated as the

THEOREM. *Let G be a finite connected graph with or without loops and multiple edges. Consider the sequence* $I^k(G)$, $k=1, 2, \cdots$. *This sequence converges if the degree of each edge is at most two; otherwise, it diverges. More precisely,*

(1) $I^k(G)=0$*-graph for* $k \geqq n$, *where n is the number of vertices in G, if G has no loops, and the degree of each vertex is* $\leqq 2$, *with the strict inequality holding for at least one vertex;*

(2) $I^k(G)=I(G)$ *for* $k \geqq 1$, *if G has no loops and the degree of each edge is* 2;

(3) *in all other cases,* n_k *and* m_k *are strictly increasing sequences for* $k \geqq n_0$ (*where* n_0 *is the number of vertices in G*).

Proof. The assertions (1) and (2) follow directly from the Cases 1 and 2 of Section 2. For the assertion (3), observe that $I^k(G)$ cannot contain vertices of degree 0 or 1 for $k \geqq n_0$, where n_0 is the number of vertices in G; and use Cases 2, 3, 4 of Section 2.

REMARKS. (1) In the assertion (2) of the Theorem, we refer to the degrees of edges (instead of the degrees of vertices as done in Section 2) in order to incorporate the exceptional graph G_1 of Figure 3. We could restate it in another form: If G is the same as G_1 of Figure 3, then $I^k(G)=I(G)$; and if G has no loops and the degree of each vertex is 2, then $I^k(G)=G$.

(2) In the case of an arbitrary graph G, some obvious modifications have to be made. Each of the connected components should be considered separately.

(3) As a simple corollary, we deduce the following result: the three statements (1) $G=I^n(G)$ for *some* n; (2) $G=I^n(G)$ for *all* n; (3) G has no loops and the degree of each vertex is 2; are equivalent for an arbitrary graph G. A different proof appears in [2].

References

1. O. Ore, Theory of graphs, Amer. Math. Soc. Colloq. Publ., 38 (1962) 21, Problem 7.

2. V. V. Menon, On the isomorphism between graphs and their adjoint graphs, Canad. Math. Bull., 8 (1965) 7–15.

RADICAL SUBRINGS OF MATRIX RINGS

GIORGIO SORANI, University of Illinois

Although nilpotent ideals are always included in the Jacobson radical of a ring, it is a rather special situation that will force the Jacobson radical itself to be nilpotent. Although special, however, it is not a rare thing, for we know it to be the case for finite dimensional algebras and rings with descending chain conditions.

In this note we give some criteria that radical subrings of matrix rings be nilpotent. We shall consistently use the following notation: T will be a commutative integral domain, T_n the ring of $n \times n$ matrices over T; for any ring R, $J(R)$ will denote its Jacobson radical.

DEFINITION: $\mathfrak{M}=\{T \mid R \subset T_n$ *implies that* $J(R)$ *is nilpotent*$\}$.

We begin with

LEMMA 1. *Let T be a commutative integral domain; suppose that T has a set of prime ideals $\{P_\alpha\}$ such that $\bigcap_\alpha P_\alpha = (0)$ and each T/P_α is finite. Then $T \in \mathfrak{M}$.*

Proof. Since P_α is a two-sided ideal of T, $(P_\alpha)_n$ is a two-sided ideal of T_n. Moreover, $T_n/(P_\alpha)_n$ is isomorphic to $(T/P_\alpha)_n$ so by our assumptions on P_α, $T_n/(P_\alpha)_n$ is isomorphic to the ring of all $n \times n$ matrices over a finite field. Let R be a subring of T_n and $J(R)$ the Jacobson radical of R. Under the natural homomorphism f_α of T_n onto $T_n/(P_\alpha)_n$, $f_\alpha(J(R))$ is a finite radical ring, hence nilpotent [1]. In fact, it is a nilpotent subring of the $n \times n$ matrices over a field; we can say that $f_\alpha(J(R))^n = (0)$, which translates into $J(R)^n \subset (P_\alpha)_n$ for every P_α. Since $\bigcap_\alpha P_\alpha = (0)$ we get $J(R)^n = (0)$ thereby proving the lemma.

COROLLARY 1. *The ring of integers is in $\mathfrak{M}$.*

COROLLARY 2. *The ring of integers of any algebraic number field is in $\mathfrak{M}$.*

We refine Lemma 1 to a sharpened form in

LEMMA 2. *Let T be a commutative integral domain having a set of prime ideals $\{P_\alpha\}$ such that $\bigcap_\alpha P_\alpha = (0)$ and each T/P_α is in $\mathfrak{M}$. Then $T \in \mathfrak{M}$.*

Proof. We proceed as before, mapping T_n on $T_n/(P_\alpha)_n$. Since $T_n/(P_\alpha)_n$ is isomorphic to $(T/P_\alpha)_n$ and $T/P_\alpha \in \mathfrak{M}$, we know that any radical subring of $T_n/(P_\alpha)_n$ is nilpotent, in fact of index of nilpotence at most n (since T/P_α is an integral domain). So if $R \subset T_n$, since $J(R)$ maps into a radical subring of $T_n/(P_\alpha)_n$, $J(R)^n \subset (P_\alpha)_n$ we see that $J(R)^n = (0)$. Thus $T \in \mathfrak{M}$.

Let $T[x]$ be the ring of polynomials in x over T. We now prove our main result.

THEOREM. *If $T \in \mathfrak{M}$ then $T[x] \in \mathfrak{M}$.*

Proof. If T is finite it is a field; for any integer $m \geqq 1$ we can find in $T[x]$ an irreducible polynomial $f_m(x)$ of degree m. Let P_m be the ideal of $T[x]$ generated by $f_m(x)$; clearly $\bigcap_m P_m = (0)$, the P_m are prime in $T[x]$ and $T[x]/P_m$ is a finite extension of the field T, so it is finite. By Lemma 1 we see that $T[x] \in \mathfrak{M}$.

Suppose then that T is infinite. If $t \in T$ then the ideal $P_t = (x-t)$ of $T[x]$ is prime, since $T[x]/(x-t)$ is isomorphic to T which is an integral domain. Moreover $\bigcap_{t \in T} P_t = (0)$, since any polynomial in this intersection would vanish on all $t \in T$, and would have an infinite number of roots in T, which is impossible if it is not zero. Summarizing, we have a set of prime ideals $\{P_t\}$ in $T[x]$ whose intersection is (0) and such that $T[x]/(x-t) \approx T \in \mathfrak{M}$. By Lemma 2, $T[x]$ is in $\mathfrak{M}$.

COROLLARY 1. *If $T \in \mathfrak{M}$ then $T[x_1, \cdots, x_r] \in \mathfrak{M}$.*

COROLLARY 2. *If Z is the ring of integers then $Z[x_1, \cdots, x_r] \in \mathfrak{M}$.*

COROLLARY 3. *If F is a finite field then $F[x_1, \cdots, x_r] \in \mathfrak{M}$.*

Note that Q, the field of rational numbers, is not in $\mathfrak{M}$ since in it, we have the local ring $\{n/m \mid n, m \text{ integers}, m \text{ odd}\}$ whose radical is $\{2n/m\}$ which certainly is not nilpotent.

This note is part of the author's thesis for the Laurea at the University of Rome.

I should like to thank Prof. I. N. Herstein for the help and encouragement given me on this work.

Reference

1. I. N. Herstein, Theory of rings. University of Chicago Press, 1951.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

ON MEASURABLE FUNCTIONS, CONTINUOUS FUNCTIONS AND SOME RELATED CONCEPTS

ELIZABETH H. YEN AND H. R. VAN DER VAART, North Carolina State University at Raleigh

I. Introduction. When comparing various authors in the field of measure theory, one observes a confusing variety of statements concerning measurable functions, continuous functions and a few related subjects. The aim of the present paper is to discuss a number of these statements within one common framework so as to resolve apparent inconsistencies and clarify interrelations between various definitions of various concepts. The need for such a comparative evaluation becomes particularly obvious in view of such statements as:

(1) A measurable function of a measurable function is measurable. ([**7**], Ch. 1, Sec. 1G; [**3**], Ch. 1, Sec. 1.2.)

(2) If ϕ is a measurable function on R, and f is measurable, it is not in general true that the composite function $\phi(f(\cdot))$ is measurable, ([**11**], Sec. 3.3).

(3) A Lebesgue measurable function of a measurable function is not necessarily measurable, ([**4**], Sec. 19, problem 3).

(4) A Borel measurable function of a μ-measurable function is μ-measurable, ([**6**], Ch. 6, Th. 6, also see [**4**], Sec. 19, problem 2, and [**1**], Sec. 13, Th. 3).

(5) An example can be given where a measurable function of a continuous function produces a nonmeasurable function, ([**8**], sec. 19).

(6) A continuous function of a (μ)-measurable function is (μ)-measurable, ([**7**], Ch. 2, Sec. 5.3; [**6**], Ch. 6, Th. 6 Cor.; [**11**], Lemma 3.3c, Sec. 3.3).

(7) A continuous (real) function is (Borel) measurable, ([**1**], Sec. 13; [**6**], Ch. 6, Th. 6, remark; [**8**], Sec. 19; [**9**], Ch. 3, Sec. 5; [**11**], Sec. 3.3, Th. 3.3.b).

(8) A continuous function (on a topological measurable space) is not in general measurable, ([**3**], Ch. 2).

The present paper does not claim the introduction of new concepts or new results. It does offer a unifying approach to a subject which by its apparent

EMBEDDING AN INTEGRAL DOMAIN USING CYCLIC ELEMENTS

B. L. McAllister, South Dakota School of Mines and Technology

It is well known that every integral domain D may be embedded in a field F. (E.g. the ring of real integers can be embedded in the field of rationals.) In the usual proof, elements of F are equivalence classes of ordered pairs (a, b) of elements of D with $b \neq 0$; the arithmetic of F is induced from that of D. We observe that the postponement of the exclusion $b \neq 0$ until late in the game permits a simple application of the theory of cyclic elements in the form due to Radó and Reichelderfer [1]. The resulting structure may shed a little light on "infinity" and "indeterminacy" in elementary mathematics.

Let P denote the set of *all* ordered pairs of elements of D, and define the relation R over P by

$$(a, b)\ R\ (c, d) \Leftrightarrow ad = bc.$$

The relation R restricted to P_1, the set of ordered pairs (a, b) with $b \neq 0$, is the usual equivalence and, in fact, R restricted to $P_2 = P - \{(0, 0)\}$ is still an equivalence. But, since transitivity fails on P itself, R does not partition P in the usual sense.

Nevertheless, R is nearly transitive on P. If $a_1, a_2, \cdots, a_n$ are members of P such that $a_1Ra_2Ra_3R \cdots Ra_nRa_1$, then for $i, j = 1, 2, 3, \cdots, n$, a_iRa_j. This property, called *cyclic transitivity* is used by Radó and Reichelderfer to decompose sets into not-quite-disjoint subsets called *cyclic elements*. For the treatment of this decomposition in full generality, see [1].

Taking a short-cut in our simple case, we may define a *cyclic element* to be either $\{(0, 0)\}$ or a subset C of P maximal with respect to the property that each two members (a, b) and (c, d) of C bear the relation R to each other. ($C - \{(0, 0)\}$ is then one of the equivalence classes into which R partitions P_2.) When D is the ring of integers, P is the set of lattice points in the plane, and a cyclic element C is the set of all points of P on some line L through the origin. Note that C degenerates if and only if the slope of L is irrational.

For each (a, b) we let $[a, b]$ denote the cyclic element to which (a, b) belongs, except that since $(0, 0)$ is in every cyclic element, we make the notation unambiguous by letting $[0, 0] = \{(0, 0)\}$. As usual, we identify $[a, 1]$ with a, and in particular $[0, 1]$ with 0. We also associate the symbol ∞ with $[1, 0]$.

Our procedure now differs very little from the usual. First we define a "pseudo-product" of members of P by members of P by the rule

$$(a, b) \cdot (c, d) = (ac, bd). \tag{1}$$

Then define $[a, b] \cdot [c, d]$ to be the smallest cyclic element C that contains all products of members of $[a, b]$ by members of $[c, d]$.

Then

$$0 \cdot \infty = [0, 1] \cdot [1, 0] = [0, 0]$$

and

$$\infty \cdot 0 = [1, 0] \cdot [0, 1] = [0, 0].$$

This is, in a sense, a "reason" for the indeterminacy of $0 \cdot \infty$ and $\infty \cdot 0$ in elementary mathematics. The only other indeterminacies of multiplication arise from using $[0, 0]$ as a factor, when the product is always again $[0, 0]$. The product $\infty \cdot \infty$ is, as it should be, "well behaved," being $[1, 0] \cdot [1, 0] = [1, 0]$.

The analogous pattern for quotients "explains" the indeterminacy of $0/0$ and ∞/∞ as well as the relatively good behavior of $1/0$ and $1/\infty$, etc.

For sums, we replace (1) by

$$(a, b) + (c, d) = (ad + bc, bd).$$

Indeterminacy, i.e. $[0, 0]$, arises only when $[0, 0]$ is used as a summand and, of course, in the case $\infty + \infty$. (Remember, here, that $[1, 0]$ is better analogous to the ∞ of the complex sphere than to the $+\infty$ of the extended real numbers.)

The identities $[0, 1]$ and $[1, 1]$ do not compose properly with $[0, 0]$, and so we do not have a field. (Moreover, neither $[1, 0]$ nor $[0, 0]$ has an additive nor a multiplicative inverse.) Of course, the field properties can be regained by finally deleting $[1, 0]$ and $[0, 0]$ after all.

Reference

1. T. Radó and P. Reichelderfer, Cyclic transitivity, Duke Math. J., 6 (1940) 474–485.

THEOREM OF GEORGE B. THOMAS, JR.

LEE YONG-JENG, Taipei, Taiwan (Now at Emory University)

Consider the sums

$$A(n) = \sum_{k=1}^{n} k^2 = \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n, \tag{1a}$$

and

$$B(n) = \sum_{k=1}^{n} 2k = (n^2 + n + \tfrac{1}{6}) - (\tfrac{1}{6}). \tag{1b}$$

If the discrete variable n is replaced by the continuous variable x in the polynomials A and B, we observe that

$$B(n) = A'(n) - A'(0), \tag{2}$$

where $A'(x)$ is the derivative of the polynomial A, at x. Observe that the summands in (1a) and (1b) are of the form

$$f(k) = k^2 \qquad \text{and} \qquad f'(k) = 2k.$$

Equation (2) also holds when the summands in Eqs. (1a, b) are replaced by k^3 and $3k^2$, or by cos k and $-\sin k$, respectively. This suggests the following

THEOREM. *Let f and F be differentiable functions on the domain of non-negative real numbers and such that*

$$f(x) = F(x) - F(x-1) \qquad \text{for } x \geqq 1. \tag{3}$$

Then

$$\sum_{k=1}^{n} f(k) = F(n) - F(0) \tag{4a}$$

and

$$\sum_{k=1}^{n} f'(k) = F'(n) - F'(0). \tag{4b}$$

Proof. Equation (4a) follows at once by summing both sides of Eq. (3). To get Eq. (4b), differentiate both sides of Eq. (3) and then sum.

Examples. We can find $F(n) = \sum_{k=1}^{n} k^3$ by applying the theorem and Eq. (1a). Here $F'(n) - F'(0) = \sum_{k=1}^{n} 3k^2 = n^3 + \frac{3}{2}n^2 + \frac{1}{2}n$ so that

$$F(n) = \frac{n^4}{4} + \frac{n^3}{2} + \frac{n^2}{4} + C_1 n + C_2,$$

where $C_1 = F'(0)$ and $C_2 = F(0) = 0$. We can evaluate C_1 from the known result $F(1) = 1$ and get $C_1 = 0$. Hence,

$$\sum_{k=1}^{n} k^3 = \tfrac{1}{4}(n^4 + 2n^3 + n^2). \tag{5}$$

As another example, consider $F(n) = \sum_{k=1}^{n} \sin k\alpha$, α constant. Applying the theorem twice, we obtain $-\alpha^2 F(n) = \sum_{k=1}^{n} -\alpha^2 \sin k\alpha = F''(n) - F''(0)$. We solve the differential equation $-\alpha^2 y = y'' + k$, and find that

$$F(n) = [\cos \tfrac{1}{2}\alpha - \cos \tfrac{1}{2}(2n+1)\alpha]/2 \sin \tfrac{1}{2}\alpha.$$

The author is indebted to the referee and Professor George B. Thomas, Jr., for their valuable suggestions.

CHANGE OF VARIABLES TEST FOR CONVERGENCE OF SERIES

LEE YONG-JENG, Taipei, Taiwan (Now at Emory University)

The integral test for convergence of series can be applied to show that

$$\sum_{n=1}^{\infty} \frac{1}{n} \quad \text{and} \quad \sum_{n=2}^{\infty} \frac{1}{n \log n} \qquad \text{diverge,} \tag{1}$$

while

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \quad \text{and} \quad \sum_{n=2}^{\infty} \frac{1}{n(\log n)^2} \quad \text{converge.} \tag{2}$$

If we look at the two series in (1) together, or at (2a, b) together, we see that the integral test for (1a) leads us to consider the question of convergence of

$$\int_1^{\infty} \frac{1}{x}\, dx, \tag{3a}$$

and for (1b) we consider

$$\int_2^{\infty} \frac{1}{x \log x}\, dx = \int_{\log 2}^{\infty} \frac{1}{u}\, du. \tag{3b}$$

The lower limits on the integrals in (3a) and (3b) are of no particular significance; it is the behavior for large values of n that determines convergence or divergence of the series, and of the related integrals.

Similarly, convergence of (2a) and (2b) is a consequence of the convergence of the following integrals:

$$\int_1^{\infty} \frac{1}{x^2}\, dx \quad \text{and} \quad \int_2^{\infty} \frac{1}{x(\log x)^2}\, dx = \int_{\log 2}^{\infty} \frac{du}{u^2}.$$

These examples suggest the following theorem.

THEOREM. *Let F be a positive, continuous, and decreasing function on the domain $x \geqq 1$. Let t be an increasing and differentiable function on a domain $x \geqq m$, where m is some positive integer, and $t(\infty) = \infty$. Then $\sum_{n=1}^{\infty} F(n)$ and $\sum_{n=m}^{\infty} F(t(n))t'(n)$ both converge or both diverge, provided $F(t(x))t'(x)$ is decreasing for $x > m$.*

Proof. With respect to convergence or divergence the first series behaves like $\int^{\infty} F(x)dx$ and the second behaves like

$$\int^{\infty} F(t(x))t'(x)dx = \int^{\infty} F(u)du.$$

Example. The series $\sum 1/(1+n^2)$ converges by comparison in the usual way with $\sum 1/n^2$. Let $t(n) = n^2$, $t'(n) = 2n$. Then

$$\sum F(t(n)) \cdot t'(n) = \sum \frac{2n}{1 + n^4},$$

so $\sum 2n/(1+n^4)$ converges by the theorem.

The author is indebted to the referee and Professor George B. Thomas, Jr., for their valuable suggestions.

CONVERGENCE TESTS BASED ON $\sqrt[n]{a_n}$

S. W. REYNER, South Dakota School of Mines and Technology

There are numerous convergence tests which are dependent on the ratio a_{n+1}/a_n, yet no common test, other than the root test, is dependent on $\sqrt[n]{a_n}$. If both a_{n+1}/a_n and $\sqrt[n]{a_n}$ have limits, these limits are the same, so one might expect to find other tests dependent on $\sqrt[n]{a_n}$. We derive convergence tests analogous to the Kummer-Jensen tests, but using $\sqrt[n]{a_n}$ in place of a_{n+1}/a_n.

In the following, let $L(i, n)=\ln(\ln \cdots (\ln n))$ (i times) with $L(1, n)=\ln n$, $L(0, n)=n$, and $L(-1, n)=1/\ln n$; let

$$P(r, n) = \prod_{i=1}^{r} L(i, n)$$

with $P(0, n)=1$ and $P(-1, n)=1/n$; let

$$S(r, n) = \sum_{i=1}^{r} L(i, n)$$

with $S(0, n)=0$; and let $c_n=\sqrt[n]{a_n}$. We suppose throughout that the numbers a_n are nonnegative. Of the following two theorems the second is a partial converse of the first.

THEOREM I. *If there exists a positive integer r and a positive integer N such that*

$$\text{(1)} \quad \frac{L(r-1, n)}{L(r, n)}\left(\cdots \left\{\frac{L(1, n)}{L(2, n)}\left[\frac{L(0, n)}{L(1, n)}(1-c_n)-1\right]-1\right\} - \cdots -1\right) > K > 1$$

for all $n>N$, then $\sum a_n$ is convergent.

Proof. First we will prove by induction on r that if

$$\text{(2)} \quad \frac{L(r-1, n)}{L(r, n)}\left(\cdots \left\{\frac{L(1, n)}{L(2, n)}\left[\frac{L(0, n)}{L(1, n)}(1-c_n)-1\right]-1\right\} - \cdots -1\right) > K(r,n)$$

for all $n>N$, then we must have

$$\text{(3)} \qquad a_n < (1 - [K(r, n)L(r, n) + S(r-1, n)]/n)^n.$$

If $r=1$, then (2) requires

$$\frac{L(0, n)}{L(1, n)}(1 - c_n) = \frac{n}{\ln n}(1 - c_n) > K(1, n),$$

$$(1 - c_n) > \frac{\ln n}{n} K(1, n),$$

$$1 - \frac{\ln n}{n} K(1, n) > c_n = \sqrt[n]{a_n}, \left[1 - \frac{\ln n}{n} K(1, n)\right]^n > a_n$$

and (3) holds ($n>N$).

Suppose (2) implies [with $r=k$, $n>N$] that (3) holds [with $r=k$, $n>N$]. Now considering (2) [with $r=k+1$, $n>N'$], multiplying both sides of (2) by $[L(k+1, n)/L(k, n)]$ followed by adding one to both sides, we have for $n>N'$

$$\frac{L(k-1, n)}{L(k, n)}\left(\cdots \left\{\frac{L(1, n)}{L(2, n)}\left[\frac{L(0, n)}{L(1, n)}(1-c_n)-1\right]-1\right\}- \cdots -1\right)$$
$$> 1+\frac{L(k+1, n)}{L(k, n)}K(k+1, n).$$

But this is merely (2) with $K(k, n)=1+[L(k+1, n)/L(k, n)]K(k+1, n)$; hence (3) holds and gives for $n>N'$

$$a_n < \left(1-\left[\left\{1+\frac{L(k+1, n)}{L(k, n)}K(k+1, n)\right\}L(k, n)+S(k-1, n)\right]\Big/ n\right)^n$$
$$= (1-[K(k+1, n)L(k+1, n)+S(k, n)]/n)^n,$$

which is exactly (3) with $r=k+1$, $N=N'$. Consequently if (2) holds, then (3) holds. In particular if $K(r, n)=K$, K a constant, and if $n>N$, then (1) implies that

$$a_n < (1-(\alpha/n))^n = [\{1-(\alpha/n)\}^{n/\alpha}]^\alpha$$
$$< \exp(-KL(r, n)-S(r-1, n)) = \frac{1}{nP(r-2, n)[L(r-1, n)]^K},$$

(where $\alpha=KL(r, n)+S(r-1, n)$) since for $\phi>0$, $(1-\phi/n)^{n/\phi}<e^{-1}$. The right hand side is the nth term of the convergent (for $K>1$) logarithmic series, so by the comparison test, $\sum a_n$ is convergent.

THEOREM II. *If there exist a positive integer r and a positive integer N such that for $n>N$*

$$\frac{L(r-1, n)}{L(r, n)}\left(\cdots \left\{\frac{L(1, n)}{L(2, n)}\left[\frac{L(0, n)}{L(1, n)}(1-c_n)-1\right]-1\right\}- \cdots -1\right)\leqq 1, \tag{4}$$

then $\sum a_n$ is divergent.

Proof. From (4) we obtain (proceeding as in the proof of Theorem I)

$$\begin{aligned} a_n &\geqq [1-(\beta/n)]^n \\ &= [\{1-(\beta/n)\}^\beta][\{1-(\beta/n)\}^{(n/\beta)-1}]^\beta \\ &> \tfrac{1}{2}e^{-\beta} = \frac{1}{2nP(r-1, n)}, \qquad \text{where } \beta = S(r, n). \end{aligned} \tag{5}$$

The strict inequality above follows since

$$\{1-(\beta/n)\}^\beta \to 1 \quad \text{and since} \quad \{1-(\beta/n)\}^{(n/\beta)-1} > e^{-1}.$$

But $1/[2nP(r-1, n)]$ is the nth term of a divergent ($K=1$) logarithmic series and therefore $\sum a_n$ is divergent by the comparison test.

COROLLARY I. *If there exist a $\theta<1$ and a positive integer N such that for $n>N$*

$$n^\theta(1 - c_n) > K > 0,$$

then $\sum a_n$ is convergent.

Proof. This is immediate from Theorem I with $r=1$ since for $n>N$, $nK/[2(\ln n)]>n^\theta$ and $(n/\ln n)(1-c_n)>2$.

Prepared in a National Science Foundation Undergraduate Science Education Program in Mathematics at Colorado State University under the direction of Professor E. R. Deal.

ON POINTS OF CONTINUITY

D. B. GOODNER, Florida State University

The theory of functions of a real variable contains many results which seem contrary to intuition. The purpose of this note is to present such a result.

THEOREM. *Let R be the real line with the usual metric, and let Y be a normed linear space with the natural (norm) metric. There is a mapping from R into Y which is continuous only on a dense set of second category and of Lebesgue measure zero.*

Proof. Let A_n be a generalized Cantor set (cf. [1] p. 88, example 4) in the closed interval $[0, 1]$ and let A_n be of Lebesgue measure $(n-1)/n$, $n=1, 2, \cdots$. It follows that each A_n is closed and is nowhere dense, that $A=A_1\cup A_2\cup \cdots$ is of first category and is of measure one, and that the complement $A'=\{x: x\in[0, 1], x\notin A\}$ of A relative to $[0, 1]$ is of second category and is of measure zero (cf. [1] p. 99, example 20). Also, A' is dense in $[0, 1]$ because it is the countable intersection of dense open sets in a complete metric space [2, p. 14]. Now let $B_1=A_1$ and let

$$B_n = \{x: x \in A_n, x \notin A_1 \cup A_2 \cup \cdots \cup A_{n-1}\} \quad \text{for } n = 2, 3, \cdots.$$

We note that $A=B_1\cup B_2\cup \cdots$ and that $B_m\cap B_n=\emptyset$ when $m\neq n$.

Let θ be the zero element in Y, let $\|y\|$ denote the norm of the element y in Y, let $\{y_n\}_{n=1}^\infty$ be a sequence of elements in Y such that $\|y_n\|=2^{-n}$, and let the mapping f be defined on $[0, 1]$ into Y by $f(x)=y_n$ if $x\in B_n$, and $f(x)=\theta$ if $x\in A'$ (cf. [1] p. 30, example 23). The mapping f is discontinuous at each point in A: if $p\in B_n$, then $\|f(p)\|=2^{-n}$ while each neighborhood of p contains a point $q\in A'$ where $\|f(q)\|=\|\theta\|=0$. The mapping f is continuous at each point in A': let $p\in A'$, let $\epsilon>0$, choose k so that $2^{-k}<\epsilon$, and then choose a neighborhood N of p that excludes $A_1\cup A_2\cup \cdots \cup A_k$; if $x\in N$, $\|f(x)-f(p)\|=\|f(x)\|<2^{-k}<\epsilon$.

Let f be as above, let $r\in R$, and let n be the unique integer such that $m\leqq r<m+1$. We define the mapping F on R into Y by $F(r)=f(r-m)$. The mapping

F is continuous only on the set $S=\{x\in R: x-m\in A' \text{ for some integer } m\} =\bigcup_m(m+A')$. Since S is everywhere dense, is of second category, and is of measure zero, F is the desired mapping.

References

1. B. R. Gelbaum and J. M. H. Olmsted, Counterexamples in Analysis, Holden-Day, San Francisco, 1964.

2. A. C. Zaanen, An Introduction to the Theory of Integration, 2nd ed., North-Holland, Amsterdam, 1961.

POWERS IN GROUPS

LADISLAUS RÉDEI, Szeged, Hungary

In semigroups, the powers of an element obey the rule

$$a^xa^y = a^{x+y} \qquad (x, y = 1, 2, \cdots).$$

Corresponding statements hold for $x, y=0, 1, 2, \cdots$ in semigroups with identity, and for $x, y=0, \pm 1, \pm 2, \cdots$ in groups. The proof for groups is usually assumed to require a breakdown into "cases," and is left by most textbooks to the reader. A complete and elegant proof may be obtained, however, by reduction to the semigroup case as follows.

For arbitrary a in the group, let (x, y) denote the assertion "$a^xa^y=a^{x+y}$ for all nonnegative integers x, y" and let (x, y, z) denote the assertion "$a^xa^ya^z=1$ for all integers x, y, z such that $x+y+z=0$." For x, y nonnegative, (x, y) and (x, y, z) are obviously equivalent. Furthermore, in a group, the equations $uvw=1$, $wuv=1$ and $w^{-1}v^{-1}u^{-1}=1$ are all equivalent, and therefore the triples (x, y, z), (z, x, y), (y, z, x), $(-z, -y, -x)$, $(-y, -x, -z)$, $(-x, -z, -y)$ are equivalent. In at least one of these triples, the first two components are nonnegative and we have the desired reduction.

MORE ON THE INFINITE PRIMES THEOREM

R. L. HEMMINGER, Vanderbilt University

Call a sequence $\{F_n\}_{n=1}^{\infty}$ of positive integers acceptable if (1) $F_n\neq F_m$ for $n\neq m$, (2) $(F_n, F_m)=1$ for $(n, m)=1$, (3) there exists a prime p for which F_p is not prime, and (4) $F_p\neq 1$ for p a prime.

The purpose of this note is to observe that *any* acceptable sequence can be used to establish the infinitude of the primes in exactly the same manner as the Fibonacci numbers were used in [1].

Let $\{F_n\}_{n=1}^{\infty}$ be an acceptable sequence and suppose that $p_1, p_2, \cdots, p_k$ are all the prime numbers. Then, by (1), (2), and (4), $F_{p_1}, F_{p_2}, \cdots, F_{p_k}$ are pairwise relatively prime integers greater than 1. Thus, since there are only k primes, each of them has only one prime factor. But this contradicts (3).

One sees immediately that, in addition to the Fibonacci numbers, the Mer-

senne and Fermat numbers form acceptable sequences since, in each case, the terms of the sequence are relatively prime.

Finally one notes that the essence of the classical proof is recaptured by considering the acceptable sequence $\{F_n\}_{n=1}^{\infty}$, defined recursively by $F_1=2$ and $F_{n+1}=F_1 \cdots F_n+1$.

Reference

1. M. Wunderlich, Another proof of the infinite primes theorem, this MONTHLY, 72 (1965) 305.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland
COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and
JOHN A. BROWN, University of Delaware

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

THE THEORY OF NUMBERS AS A REQUIRED COURSE IN THE COLLEGE CURRICULUM FOR MAJORS

I. A. BARNETT, University of Cincinnati, Ohio University

The role of the theory of numbers in the development of mathematics is well known. Such eminent mathematicians as Hilbert, Hardy, and Dickson have expressed themselves vigorously on the benefits to be derived from a study of the subject. In his familiar essay, "A Mathematician's Apology," which first appeared nearly forty years ago, Hardy writes:

> "The elementary theory of numbers should be one of the very best subjects for early mathematical instruction. It demands very little previous knowledge; its subject matter is tangible and familiar; the processes of reasoning which it employs are simple, general and few; and it is unique among the mathematical sciences in its appeal to natural human curiosity. A month's intelligent instruction in the theory of numbers ought to be twice as instructive, twice as useful, and at least ten times as entertaining as the same amount of Calculus for Engineers."

The purpose of this brief article is to expand on Hardy's statement and to show why the subject matter of the theory of numbers should be required not only of all mathematics majors, but also of all prospective teachers of elementary-school arithmetic as well as teachers of high-school algebra and geometry.

There are, of course, many colleges that offer a course in the Theory of Numbers; and it is usually elected as a sort of "filler" course. Those in charge of programs for mathematics majors have overlooked one of the most effective disciplines for the training of students. In the opinion of the writer, to discover mathematical talent, there is no better course in elementary mathematics than number theory. Any student who can work the exercises in a modern text in

number theory should be encouraged to pursue a mathematical career. Each problem requires a type of reasoning and insight that is hard to match in any other subject. Furthermore, the student may check his results and thus obtain confidence which he could not find in routine problems of the calculus. Not only majors can profit from such study. For the prospective teacher, notions of divisibility and other concepts of number theory will serve as a basis for a better understanding of arithmetic and algebra.

To be sure, a chapter or two on number theory appears in most textbooks on abstract algebra, but the main purpose of such material is to give the student special instances and illustrations to clarify the definitions. This material, however, does not suffice to initiate the student to the real beauty of the subject nor to develop in him the insight that is so important in the further understanding of mathematics.

A first course in number theory should not be given as a unified mathematical discipline. Nor should it be presented as a series of disconnected puzzle problems. The instructor should strike a middle course; he should give the student an historical perspective into the importance that many great mathematicians have attached to the subject. The mathematics major will be challenged by the deceptive simplicity of the problems; his curiosity, a trait so basic to the study of mathematics, will be aroused.

To point out a few places which serve to emphasize the beauty and the nature of the material, one may begin with the definition of the greatest common divisor (g.c.d.). The definition may be given in two ways:

A. *d is the* g.c.d. *of two integers, a and b, if d is a divisor of a and of b; and every divisor of a and b is a divisor of d.*

B. *d is the* g.c.d. *of a and b, if there are integers α and β such that $a=\alpha d$, $b=\beta d$, where α and β are relatively prime.*

Prove that A implies B and B implies A.

In practically all arguments involving divisibility, the Fundamental Theorem of arithmetic is used. This Theorem may be stated in two equivalent forms:

A. *Every integer may be factored uniquely as a product of primes.*

B. *If the product of two integers is divisible by a third, and one of the factors of the product is relatively prime to the third, then the other factor is divisible by the third integer.*

The proof of the equivalence of these two definitions is an instructive exercise. It is rather remarkable that such an elementary fact as the g.c.d. of ka and kb equals k times the g.c.d. of a and b, requires for its proof the Euclidean Algorithm or something equivalent to it.

The student majoring in mathematics should find satisfaction in learning about the result concerning the sum of the divisors of an integer, its application

to perfect numbers, and the connection of perfect numbers with Mersenne primes. He will certainly be astonished to learn about Fermat primes and their relation to the Euclidean constructions of regular polygons.

The properties of Pythagorean triangles and the fact that their areas cannot be perfect squares will give him an introduction to certain types of Diophantine equations which have no solutions. Here he will learn about Fermat's method of infinite descent. The depth of the properties of quadratic residues and the quadratic reciprocity theorem will give him a respect for the inventiveness of the human mind. When he sees how irregularly the primes are distributed, he will appreciate how the Prime Number Theorem brings order out of chaos.

Finally, he will be impressed with the simplicity and beauty of Euclid's proof concerning the infinitude of primes. The climax of Euclid's discovery comes with Dirichlet's result that every arithmetical progression $an+b$, where a and b are relatively prime, contains an infinitude of primes. Not all of these results need be proved in a first course in number theory, but they should certainly be stated and illustrated.

One could go on citing many other remarkable results, as well as problems which still await solution. One of the best reasons for requiring a course in number theory for every major is that it will give the beginner an opportunity to start research at an early stage. The student can make up conjectures which he can prove to be true or false. He will find that the amateur as well as the professional mathematician can discover new results. One need only cite the example of how a young nineteenth-century Italian boy found a new pair of amicable numbers. This pair, 220 and 1184, which is relatively small in magnitude, had been overlooked by the professional mathematicians.

This short outline should show that there is a veritable gold mine of material which could serve as an excellent introduction to mathematics for young students, and which will arouse their curiosity and be a challenge to their mental powers.

THE INTERCOLLEGIATE MATHEMATICS LEAGUE

STANLEY RABINOWITZ, Polytechnic Institute of Brooklyn, vice president ICML

During the 1965–66 school year, a group of college students in New York City banded together to form the Intercollegiate Mathematics League (ICML). The purpose of the League is to coordinate mathematical competitions among its member schools and by so doing to stimulate the students of these schools with the interesting problems proposed.

The problems used are selected by the Faculty Advisor, presently Professor Salkind of the Polytechnic Institute of Brooklyn, from those contributed by volunteers from the faculty of the participating colleges. The contests are an hour long and are taken by teams from each member school. A team consists of from five to ten undergraduates and the team's score is the sum of their top five scores. Both individual and team prizes are awarded.

Because of a late start, only three contests were held during the 1965–66 year. They are listed in the appendix. For this year, the winning school was Cooper Union. Second and third places went to Polytechnic Institute of Brooklyn and Adelphi University, respectively. The highest individual scorer was Harry Ploss of Cooper Union and second and third place went to Stanley Rabinowitz and Richard Lary both of Polytechnic Institute of Brooklyn. The schools received trophies for their achievement and the individuals were given books on mathematics.

The schools which are presently in the League are Adelphi University, Brooklyn College, The Cooper Union, Iona College, Newark College of Engineering, New York Institute of Technology, New York University (at Washington Square and University Heights), Polytechnic Institute of Brooklyn, Pratt Institute, Queens College, and St. John's University. Other schools in or around the Metropolitan New York area which are interested may join by contacting the president of the League, Mark Brody, at Cooper Union or the author.

Appendix. The following are the first three contests held by the ICML.

MEET I

1. Find all the positive values of a for which $\sum_{n=1}^{\infty} (a^n n!)/n^n$ converges.

2. Let a be a given positive number. For each positive number, b, the graph of $x^2/a^2+y^2/b^2=1$ is an ellipse (or a circle). Let the line with equation $x=x_1$, $0<x_1<a$, intersect each of these graphs at two points. Prove that the tangents to the graphs at the intersection points all meet the x-axis in the same point.

3. Prove that if r_1 and r_2 are the roots of $x^2-6x+1=0$, then $r_1^n+r_2^n$, n a natural number, is an integer.

4. Prove that triangles having vertices at lattice points in the xy-plane cannot be equilateral. (A lattice point is one with integer coordinates.)

5. Let $x_1+x_2+\cdots+x_n=s$ where $x_i>0$, $i=1, 2, \cdots n$, and s is fixed. Prove (a) If $p=x_1x_2\cdots x_n$, then $p(\text{maximum})=s^n/n^n$.

(b) $\sqrt[n]{(x_1x_2\cdots x_n)}\leqq(x_1+x_2+\cdots+x_n)/n$, that is, the positive geometric mean of n positive numbers can never exceed their arithmetic mean.

MEET II

1. Prove, by the use of "Taylor's Theorem with Remainder" that any curve with equation $y=f(x)$, having an inflection point at $x=x_0$, and for which the second derivative is continuous, crosses the tangent line at x_0. [Define an inflection point as one where the second derivative changes sign.]

2. Find the minimum of the function F where $F(\lambda)=\int_0^1[x^2-(x+\lambda)]^2dx$.

3. Find the set S of points $P(x, y)$ such that P is equidistant from the x-axis and the circle with center at the origin and radius a.

4. (a) Find the points x_1 and x_2 so that the formula $\int_0^1 f(x)dx=f(x_1)+f'(x_2)$ yields exact results for polynomials of degree $\leqq 2$. (You may assume f possesses

the properties you deem necessary for this problem.) (b) Determine the error (in absolute value) in using this formula for any third-degree polynomial with leading coefficient 1.

5. Define A_n recursively by $A_{n+1}=A_n^2-A_n+1$ with $A_1=2$. Prove that each A_i is relatively prime to all other A_j, $i\neq j$, $i, j=1, 2, 3, \cdots$.

Meet III

1. Find the $\lim_{x\to\infty} [x]/x$, where $[x]$ is the greatest integer less than or equal to x. Show all work.

2. Prove or disprove that you can add a constant to the variable of integration if you subtract it from the limits of integration.

3. Derive the general term $u_n(x)$ of the series for

$$\frac{1}{1-x}\ln\frac{1}{1-x}\cdot$$

4. Prove without using l'Hôpital's Rule that $I_1\approx I_2$ when s is near 1, where

$$I_1=\int_0^1\frac{dx}{(1+x)^s}\qquad\text{and}\qquad I_2=\int_0^1\frac{dx}{(1+x)}\cdot$$

5. Let f be a family of lines such that for each line, OP is the x-intercept and OQ is the y-intercept and segment PQ is constant, that is, $PQ=k\neq 0$, k constant. Find the equation of the envelope of f.

Definition. The envelope of a family f of ∞^1 curves is the curve (or curves) such that, at each of its points, it is tangent to a member of the given family f.

A MATHEMATICS COMPETITION IN CALIFORNIA

Brendan Kneale, F.S.C., St. Mary's College

In a recent report (see [1]) J. R. Smart described a program in Wisconsin which calls for sending five problem-sets each year to more than 500 high schools and selecting the top fifty or so students to come to the State University for a special day-long session. It may be of interest to compare that program with a similar but differently oriented one in California.

Six years ago the Department of Mathematics of St. Mary's College sent to thirty high schools in the State a set of ten problems for each month of the school year, followed up with correction of the papers mailed in, score sheets, and solutions. The purpose was chiefly to upgrade performance in local and national contests which involved problem-solving. The selected schools, which were mostly private or diocesan, had performed poorly in past contests. The next year forty public high schools from two Bay Area counties were invited to join the program, and a contest on the campus of the College was announced for those students who performed well. The response in that and subsequent years was most heartening. During the third year NSF gave further impetus to

the work by supporting the main part of the program (though not the final contest, support for which was against Foundation policy). By 1966 the overall program has grown considerably and most of the "bugs" seem to have been removed.

In its present form the program has the following steps:

A. At the start of the school year letters of invitation are sent to all junior and senior high schools in the San Francisco-Oakland area and to selected schools in nearby counties. (We have previously cleared the program with school system superintendents.) These letters read in substance:

> Once again you and your students are invited to participate in the Problem-Solving Program in Mathematics sponsored by Saint Mary's College and the California Mathematics Council Northern Section.
>
> The program provides high school and junior high school students with an opportunity to solve problems of some difficulty not just once a year at the time of a mathematics contest but throughout the year. The aims envisaged by the sponsors are: 1) To arouse mathematical interest, especially in key areas of problem-solving; 2) To discover mathematical talent that might otherwise remain hidden; 3) To help students keep widely separated parts of mathematics freshly in mind; 4) To direct abler students to advanced techniques, useful notations, and to habits of generalization and abstraction; 5) To upgrade performance in homework, examinations, and contests; 6) To help teachers continue developing their own skills.
>
> Two separate problem-solving programs will be conducted simultaneously, one for students of grades 10–12 which will be subtitled the senior high school division, and one for students of grades 7–9, subtitled the junior high school division. For each division there will be four problem-sets sent during the year on the dates listed in the enclosed schedule. One month after these dates solutions will be due at the College. At that time solution-sheets will be sent to the participating schools so that students may check their results while the problems are still fresh in mind. During the following two weeks the problems will be corrected, points will be assigned to the students, a report of the students' rating will be made, and the next problem-set will be sent to the schools. The pupils themselves may mail in their individual solutions or the instructors may prefer to collect the papers and mail them all in together.
>
> This program is capped with an invitational contest to be held May 7, 1966, on the campus of the College. Eligibility will be determined as follows: For each problem-set the maximum possible number of points is 100, so that the total possible number for the year is 400; only students who during the course of the year have earned a total of at least 100 points in the ratings will be allowed to participate in the contest. Further information will be sent prior to the day of the contest.

B. As in former years there were requests from school instructors for guidelines on how much help the students should receive in trying to solve the various problems. The reply that we gave this year was the following: Students are expected to learn by doing so that it is hoped that the work mailed in represents personal effort. Success in the final contest is evidence of this effort. On the other hand, the program is supposed to lead to new ideas for the students, hence whatever concepts and techniques are unknown to the students but are needed for *understanding* the problems should be taught to them, yet the students ought not to be deprived of the chance to make discoveries of such concepts and techniques for themselves.

C. In the spring letters are sent to local companies requesting that they make prize-donations in the form of books, slide rules, and other aids. These,

together with cash support from the College and the California Mathematics Council, result in an impressive array of prizes—this year containing twenty-three for the senior division and thirty-four for the junior division.

D. For the final contest in May we mailed out official eligibility lists and instructions to the schools who had students with 100 or more points, that is, to twenty-nine senior and fifty-one junior high schools. Of the 105 students from the former and the 172 from the latter some could not come to the contest because of local activity-conflicts, but the final count was very good: eighty-five and 160 in the respective divisions. Also some forty teachers attended, many of them generously assisting with the work of proctoring, correcting, and tallying.

Unlike the Wisconsin program described in [1] we are not engaged primarily in a "search for mathematical talent." Rather we design our efforts mainly to develop and encourage such talent, and to a certain extent we include the school teachers in this effort. There is evidence that many schools, classrooms, and teachers of whom we never hear take full advantage of the problems and solutions which we mail out.

Sample Problem-Set. Senior Division, April, 1966.

(1) Find the matrices with real elements which are the square root of

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

(2) A trapezoid has parallel bases of 16″ and 10″ and equal sides of 12″. A line is drawn from one end of the 16″ base to the opposite 12″ side so as to divide the trapezoid into two figures of equal area. If the division line cuts the opposite side a distance x from the end of the base of length 10″, what is x?

(3) The definition of e, the base of the natural logarithm system, is $e=\lim_{x\to 0}(1+x)^{1/x}$. On the basis of this definition find $\lim_{z\to\infty}(1+5/z)^{10z}$.

(4) Nine cards are dealt face down. Somebody looks at one of the cards and places it on top of the pack of nine cards. These are then placed under the deck, face down. The dealer now makes four piles of cards as follows. He turns cards up counting 10, 9, 8, 7, . . . and stops if one of the numbers he calls coincides with the card. If there is no coincidence by the time he reaches 1, he puts a card face down on this pile. In this manner he builds up four piles. He then adds the sum of the numbers appearing on the four piles and counts out that many cards. The last card he turns up is the card that was originally selected and put on top of the pile of nine cards. Explain why this trick works.

(5) Draw the graph of $y=|5-|5-|x|||$.

(6) There are n keys and m locks, with $n\geqq m$. One key and only one opens each lock. What is the average number of trials required to open the m locks?

(7) The roots of the equation $x^2-3x-1=0$ are $r=(3+\sqrt{13})/2$ and $s=(3-\sqrt{13})/2$. Prove that, for n a positive integer, r^n+s^n is always an integer.

(8) Work out a formula f, in the form of a polynomial of least degree, which has the following values: $f(1)=2$; $f(2)=7$; $f(3)=16$; $f(4)=30$; $f(5)=50$; $f(6)=77$; $f(7)=112$.

(9) Solve for x to three decimal places: $2^x=2^{-x}-4$.

(10) Determine whether the equation $x^2-25x+163=0$ has integral solutions in any number base.

Sample Problem-Set. Junior Division, April, 1966.

(1) Two squares, each of side s, are so placed that their common area is a regular octagon. What is the area of this regular octagon?

(2) Express 17/25 as the sum of three different fractions each having 1 as numerator and an integer as denominator.

(3) Given the following table, determine the manner in which it is built up:

1										
1	1	1								
1	2	3	2	1						
1	3	6	7	6	3	1				
1	4	10	16	19	16	10	4	1		
1	5	15	30	45	51	45	30	15	5	1

(4) Given two fractions 3/5 and 5/4 in base 7. Find their sum performing all operations in base 7. (In other words, put the fractions into a form so that they have a common denominator, and then add.)

(5) Find the set of positive integers whose sum is 20 and whose product is a maximum.

(6) Two wheels of radii 2′ 10″ and 1′ 14″ respectively are rotating in contact with each other. Let the point of contact at the start be marked on each wheel. What is the smallest number of revolutions of each wheel before the original point of contact becomes the point of contact again?

(7) What integer multiplied by 52631578947368421 will give all 9's?

(8) Given objects of four colors: red (R), blue (B), green (G), and white (W). These are to be placed in a row in such a way that (1) no two objects of the same color are together; (2) two successive objects of given colors may not be repeated in the sequence. Thus, once RW is used, this combination may not be used again. How many objects may be put in a row according to these rules? (Show your arrangement.)

(9) A student has four science books, three literature books, and two mathematics books. He wishes to keep all books of the same kind together on his shelf. In how many ways can the books be arranged?

(10) Tom Jones and Bill Smith set out to create their own thermometers. Jones calls the freezing point of water on his scale 40 degrees, while Smith calls

his freezing point 25 degrees. Jones makes the boiling point of water 280 degrees and Smith his boiling point of water 125 degrees. What temperature on the Smith scale is equivalent to 97 degrees on the Jones scale (to one-tenth of a degree)?

Reference

1. J. R. Smart, Searching for mathematical talent in Wisconsin, II, this MONTHLY, 73 (1966) 401–406.

NCTM FILM SERIES

MATHEMATICS FOR ELEMENTARY SCHOOL TEACHERS

After an investment of three years' hard work and nearly $300,000, the NCTM project Films in Mathematics for Elementary School Teachers now has tangible results to show for these efforts. A series of ten in-service teacher education films has been completed. They are 16mm color, sound films, each approximately half an hour long. Up to this point these films may sound like "just another series." This is far from true. The films are different because they were produced differently. Media experts, mathematicians, and mathematics educators were joined together in a working team from the very inception of the project. Somewhat as a "shakedown cruise," a series of five pilot films was produced. Critiques of these were gathered from many informed persons. The films served their purpose and were never released for general classroom use. From the ideas and experiences gained in this initial project, plans for the present series took shape. The team approach to producing these films is evident throughout. They are not stand-up blackboard lectures; appropriate teaching aids were specially built and are skillfully used; good teaching techniques were practiced as well as preached; a talented educator, with wide experience in the medium, plays the role of teacher. Mathematically precise and accurate content, sound teaching techniques, and technical finesse have been combined in this series of films in mathematics for elementary school teachers. A 224-page correlated text has been published to accompany the films.

The National Science Foundation has provided financial assistance in the amount of $276,000. In addition, approximately $40,000 of NCTM members' dues has been invested in the project.

BRIEF COMMENT

News of the "New Math," GEORGE WEBER, *Bulletin of the Council for Basic Education*, 10 (1966) No. 9, 1–4.

A review of some recent comments by a variety of people on topics in school mathematics is given here. E. G. Begle's plans to convene a group to design a new sequential curriculum for the schools in grades 7 through 12, as reviewed in this section earlier, is mentioned. The concern of a number of people, Bernard

Friedman and Max Beberman among them, for the loss of computational skills which appears to accompany the introduction of "new math" is brought up. Some remarks of Frank B. Allen's and of Howard Fehr's are also mentioned.

SMSG The Making of a Curriculum, WILLIAM WOOTON, Yale University Press, New Haven, Connecticut, 1965, x+182 pp.

It is not the intention of this section of the MONTHLY to review books; however, since it is likely that this book will not be reviewed in the section on book reviews, it was thought desirable at least to call attention to its existence here. "This brief book provides a description of the origin and activities of the School Mathematics Study Group (SMSG) over the first four years of its existence. The chief purpose of this book is to describe the activities and *modus operandi* of SMSG. The book is addressed to any person, mathematically trained or otherwise, with a curiosity about how a group such as SMSG undertakes to influence the curriculum of the schools of the United States." There is an extensive appendix which names all the people who were involved with SMSG in the years 1958 to 1962.

Games and Programed Instruction, LAYMAN E. ALLEN, *Programed Instruction*, V (1966), No. 6, 9–11.

The claim is made that "(1) appropriate kinds of games can themselves be a rather sophisticated and complex mode of programed instruction, and (2) that learning programs can be constructed so that they incorporate important features of games—hopefully, the features that make games self-motivating." Two games are mentioned to substantiate the claim. The first is WFF'N PROOF, a game—or better a sequence of successively more difficult games—based on the statement calculus. WFF'N PROOF is mentioned but briefly, and then the second game EQUATIONS is described in considerable detail. EQUATIONS is a game in which a randomly determined set of elements—numbers and signs of operation—must be arranged by the players to form a correct equation, subject to a rather complex set of rules. The article is reprinted from *The Arithmetic Teacher*.

Topics: The Case of the Angry Mathematician, HARRY SCHWARTZ, *The New York Times*, April 23, 1966, p. 30.

This brief and amusing editorial was prompted by the appointment of Nachman Bench by Mayor Lindsay of New York City. It makes essentially two points. The first is that by and large mathematics gets much poorer reporting in the daily press than do the humanities. The second is that the computer is a "moron machine and it can do nothing by itself." The latter point, though a familiar matter to any mathematician, is still a bit of a mystery to non-mathematicians. Its emphasis here is good to see.

Comment on Review of "Mathematics and the High School Curriculum." F. V. Pohle, Adelphi University. (See Monthly, April 1966, Part I, p. 409.)

The Review asserts that the Report "appears to be largely an apologia for and an expansion of the views of Professor Morris Kline . . . ". Appearances can be deceiving. The objective of the Report was clearly stated in the first sentence: "The following comments have been written to supply background information on a subject which cannot be discussed without a substantial amount of preliminary reading."

The material was intended for high school teachers, parents, and administrators interested in the teaching of mathematics in our high schools and was presented largely through quotations from many authorities: Courant, Dieudonné, Stone, von Neumann, Yang, Weyl, Rosenbloom, Kline, Beberman, Biot, Stoker, Diliberto, and Moise. In addition, reference was made to SMSG texts.

The Review did not note comments with respect to three important references: (1) The Cambridge Report (Goals for School Mathematics); (2) "On the Mathematics Curriculum of the High School" which appeared in this Monthly (vol. 69, No. 3, March 1962, pp. 189–193); (3) recommendations made in 1964 by Dr. John W. Gardner, now a member of President Johnson's Cabinet but written at the time when Dr. Gardner was President of the Carnegie Foundation for the Advancement of Teaching. Since precisely such recommendations have been made recently at the University of California at Berkeley and have been adopted in modified form at Yale University, the recommendations may be worth further consideration.

The following 11 articles are taken from the Proceedings of the Preliminary Meeting on College Level Mathematics Education under the Auspices of the U. S.-Japan Program on Scientific Cooperation, Katada, 1964, *Japan Society for the Promotion of Science*, 1965.

Problems of College Level Mathematics Education in Japan, Shokichi Iyanaga, pp. 1–8.

This article describes the educational system in Japan, limiting itself to the college and secondary school curricula. The description of the high school curriculum is rather more detailed, since the college curriculum is described in considerably greater detail in other articles of this booklet. After the description of the curricula, the following problems are stated: "(I) How to improve our high school mathematics? (II) Where should abstract ideas be introduced? (III) How about the logical foundations of elementary concepts? (IV) How about classical subjects such as differential geometry in R^3 or projective geometry? (V) How about numerical analysis and computer mathematics?"

Analysis for the College Level Mathematics of Japan, KOSAKU YOSIDA, pp. 9–16.

"One of the salient features of the curriculum of the mathematics education in Japan is that we teach in high schools the differential and integral calculus of functions of one independent variable, including exponential and logarithmic functions. Of course, it is taught on the ground of very intuitive notion of limit concerning real numbers." The article goes on to describe the college curriculum for the first three semesters at the College of General Education of the University of Tokyo which curriculum is fairly typical. Thereafter the curriculum for more advanced analysis courses is also given. Finally two alternative plans for revision of the basic three semester curriculum are given.

Algebra and Geometry for Science and Engineering Students, YUKIYOSI KAWADA, pp. 17–21.

Students described in the title generally take either one or two years of courses in algebra and geometry at the rate of two hours per week, the courses comprising basic ideas in analytic geometry and linear algebra. It is assumed that the students have had at the high school level the following: "polynomials, rational functions, algebraic equations of degree two, permutations and combinations, elementary geometry, plane coordinate geometry, conic sections (in canonical forms), notion of vector in plane, complex numbers, introduction of coordinates in space." Assuming this background there are presented two course outlines for the presentation of the following three main topics: "(1) Analytic geometry in space with the use of vector notation, (2) Linear algebra in n-dimensions, (3) Some concepts from abstract algebra (such as abstract groups)."

Probability and Statistics for the First Two Years of University, SHIGERU FURUYA, pp. 22–27.

An outline is given which presents pertinent information concerning the presentation of courses in probability and statistics at thirteen Japanese universities. An outline is then given for a typical course in probability and statistics, and a further outline is given for a course in statistics alone as presented at the University of Tokyo. Finally an outline for a proposed course in probability and statistics (without measure theory) is given.

Where to Place Abstract Ideas, YUKIO MIMURA, pp. 28–29.

This article deals with the problem of where in the college curriculum to place ideas associated with foundations and with the rigorous development of the real numbers. The students in Japanese colleges arrive with a background in calculus from their secondary school education, albeit a calculus taught from an intuitive standpoint. The pressure on progressing rapidly with the concepts of the calculus precludes any extensive background on foundational topics. It is

thus Professor Mimura's proposal to introduce a course sometime during the first two years that will deal separately with "the fundamental notions of modern mathematics."

College Level Mathematics Education and Computers, HIROSHI NOGUCHI, pp. 30–34.

This article proposes that in the future, when high speed computing machines may no longer be sequential, but may trend toward parallel operation, the disciplines of numerical analysis and high speed computation may be quite different. It also makes an interesting application of category theory to flow diagrams. Finally, it proposes "1) For mathematics students, start courses on algebra and general topology similar to those recommended by CUPM (green, pp. 51, 61) as early as possible (at best, within the first two years) including a little of high speed computation. 2) For engineers and physicists a course on elements of modern mathematics such as algebraic structures of CUPM (yellow without C, p. 38) and some of general topology (green, p. 61) have to be provided as elective courses at least."

Mathematics Education for Other Sciences, YASUO AKIZUKI, pp. 35–38.

"Some physicists well recognize the effectiveness of modern mathematical thinking . . . but they fear that, if their students learn mathematics from the very beginning in too abstract a way, then they might not grasp its relationship to physical matter, since the intuitional image would be hidden under the nice formulation." "Most of the liberal scientists and economists desire mathematics education to be modernized and prefer the ideas and meaning of concepts to calculating techniques." In attempting to deal with the needs of other sciences for mathematics Professor Akizuki urges that abstract concepts be illustrated insofar as possible with as many examples as possible—examples chosen from the physical world. He stresses the importance of geometrical intuition and of the global view of mathematics, not merely the appreciation of the logical development of the subject.

Why and how C.U.P. began, E. J. MCSHANE, pp. 39–48.

"Let me begin by giving a short table of contents of my talk. First I shall indicate the background of the unsatisfactory situation existing some twenty or thirty years ago. Then I shall outline the ideas underlying the early work of the Committee on the Undergraduate Program of the Mathematical Association of America, and describe some of the work done by that committee in its first six years. In this I shall try to point out some mistakes." True to his promise the article follows the course laid out and as a bonus at the end there are included some of the author's own opinions on some of the matters broached in the earlier part. The style of the article is infused with the author's usual charm and ends with, "I feel that it is important to remember in all efforts at

improving the teaching of mathematics, that we are teaching human beings, and that what we are teaching them is a human activity with uses and with beauty and with surprises."

Organizing Nationally for Improvement of College Mathematics, W. L. DUREN, JR., pp. 49–59.

This article deals with the more recent history of the Committee on the Undergraduate Program (CUPM) and with some of the problems this committee will face in the future. It also deals with the broader problem of organizing such a committee to reorganize the teaching of mathematics at the college level. "Our survey of the factors which press us to attempt a nationally organized revision of college mathematics and the forces resisting any such effort gives an idea of the problems which confront those of us who have undertaken some responsibility for it. Perhaps our greatest problem is not how to overcome the resistance in order to bring our plans to realization, but how to exercise wisdom, self-discipline and restraint, and how to retain our respect for the mathematics we have inherited from the past and preserve it. For our times are times of upheaval and change, not only in general society, but in some ways even more radically in our small world of mathematics and mathematics education."

Activity and Motivation in Mathematics, EDWIN E. MOISE, pp. 60–66.

"If we look at a four-year college program, as a whole, it is very natural to judge it by the quantity of mathematical knowledge that it appears to contain, and by the speed with which it reaches the concepts and methods of contemporary mathematics. Surely these are important criteria. I believe, however, that there are other criteria which it is easy to neglect, because they are harder to describe and to apply. In the first place, I do not believe the effect of an educational program can safely be judged merely by the quantity of knowledge which the student can demonstrate on an examination. It seems to me that this sort of simple book keeping is not just slightly inaccurate but grossly so." Much of the rest of the article goes on to document these ideas via a description of the methods of R. L. Moore, the notion of genetic principle, as well as other examples. "I believe that there are no easy answers to the questions that I have been raising. The fundamental task of a mathematician who teaches is to convey to his students not only what mathematicians know, but also what they do, and how, and why. This is a problem in full and honest self-revelation. And (as Raymond Chandler has remarked in another connection) *honesty is an art.*"

New Notions for Algebra Courses in College, SAUNDERS MACLANE, pp. 67–74.

"A fundamental reform and improvement in college courses in algebra began in 1930 under the slogan 'Modern Algebra.' A second revolution in algebra is now underway: a revolution which emphasizes the importance of studying the 'morphisms' of each type of algebraic system. These ideas entered algebra by way of its application to algebraic topology and geometry; they are formally

expressed by means of the notions of category and functor; these formal expressions have led to many new applications. It is now clear that these applications can drastically modify the style of future courses in algebra." The remainder of the article documents these notions in technical detail by describing formally what a category and a functor are, and by giving a number of examples.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J. 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to M. S. Klamkin, Ford Scientific Laboratory, P.O. Box 2053, Dearborn, Mich. 48121. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before March 31, 1967.

E 1925. *Proposed by W. J. Blundon, Memorial University of Newfoundland.*

Let r and R denote the inradius and circumradius, and define $k=r/R$, so that $0<k\leqq\frac{1}{2}$. Prove that, for every triangle ABC,

$$(1)\quad -1+4k-k^2\leqq \cos A\cos B+\cos B\cos C+\cos C\cos A\leqq k+k^2,$$

$$(2)\quad -1+3k-\tfrac{3}{2}k^2\leqq \cos A\cos B\cos C\leqq \tfrac{1}{2}k^2,$$

with equality only for equilateral triangles.

E 1926. *Proposed by L. D. Yarbrough, Harvard Computing Center*

Express in terms of N and b the sum of the digits of the integer N as written in radix b notation. (This is a generalization of the rule of "casting out nines," and for $b=2$ the formula yields the number of 1's in the binary representation of N, which is a measure of the multiplication speed of certain digital computers.)

E 1927. *Proposed by Daniel W. Burns, Woodstock, Ill.*

Given any integer $m(>6)$ which possesses primitive roots, let r be a primi-

tive root of m. Put $a_j \equiv r^j \pmod{m}$, where a_j is the least positive residue. Now form

$$F(x) = \sum_{j=1}^{\phi(m)} a_j x^{\phi(m)-j}.$$

Prove that $F(x)$ is the product of $t+1$ polynomials

$$F(x) = f_1(x) \cdot f_2(x) \cdots f_t(x) \cdot g(x),$$

where every coefficient of $f_i(x)$ is 1, and the sum of the coefficients of $g(x)$ is m. Here t is the number of distinct prime divisors of $\frac{1}{2}\phi(m)$. The result is still true if the coefficients of $F(x)$ are permuted cyclically. (The result is true for $m=5$ and, trivially, for $m=3, 4, 6$.)

E 1928. *Proposed by D. Ž. Djoković, University of Belgrade, Yugoslavia*

Let $A=(a_{ij})$ be an $n \times n$ matrix such that $a_{ii}=a_i$, $i=1, 2, \cdots, n$; $a_{i,i+1}=b_i$, $i=1, 2, \cdots, n-1$; $a_{ij}=0$ otherwise. Let M be the minor of det A obtained by deleting the rows $i_1, i_2, \cdots, i_k$ $(1 \leqq i_1 < i_2 < \cdots < i_k \leqq n)$ and the columns $j_1, j_2, \cdots, j_k$ $(1 \leqq j_1 < j_2 < \cdots < j_k \leqq n)$. Prove that

$$\begin{aligned} M = {} & (a_1 a_2 \cdots a_{j_1-1})(b_{j_1} b_{j_1+1} \cdots b_{i_1-1})(a_{i_1+1} a_{i_1+2} \cdots a_{j_2-1})(b_{j_2} b_{j_2+1} \cdots b_{i_2-1}) \\ & \cdot (a_{i_2+1} a_{i_2+2} \cdots a_{j_3-1}) \cdots (b_{j_k} b_{j_k+1} \cdots b_{i_k-1})(a_{i_k+1} a_{i_k+2} \cdots a_n) \end{aligned}$$

if $1 \leqq j_1 \leqq i_1 < j_2 \leqq i_2 < j_3 \leqq i_3 < \cdots < j_k \leqq i_k \leqq n$; $M=0$ otherwise. We take $(a_r a_{r+1} \cdots a_s)=1$ whenever $s<r$.

E 1929. *Proposed by D. M. Bloom, Brooklyn College*

Each of the n teams of a baseball league plays each other team exactly r times during the season. (There are no tied games.) Let $\|a_{ij}\|$ $(i=1, \cdots, n;$ $j=1, 2)$ be the $n \times 2$ matrix representing the final won-and-lost records of the teams, arranged in the usual order. Then it is easily seen that

(1) $$a_{i1} + a_{i2} = r(n-1) \qquad (\text{all } i),$$

(2) $$\sum_{i=1}^{n} a_{i1} = \sum_{i=1}^{n} a_{i2},$$

(3) $$\text{if } i < j, \text{ then } a_{i1} \geqq a_{j1} \text{ and } a_{i2} \leqq a_{j2},$$

(4) $$\sum_{i=1}^{k} a_{i2} \geqq rk(k-1)/2 \qquad (k = 1, \cdots, n).$$

Prove that any $n \times 2$ matrix $\|a_{ij}\|$ satisfying conditions (1) through (4) represents a possible outcome of the baseball season.

E 1930. *Proposed by Simeon Reich, Haifa, Israel*

Let a, b, c be the sides of an acute triangle, r its inradius, and r_a, r_b, r_c its exradii. Deduce:

$$(1) \qquad \left(\sum a\right)^3 \leqq 5\sum a^2b - 3abc,$$

$$(2) \qquad 9r\left(\sum r_a\right)^2 + 9r^3 \geqq 32r_ar_br_c - 14r^2\sum r_a,$$

with equality if and only if the triangle is equilateral.

E 1931. *Proposed by D. W. Burns, Woodstock, Ill.*

Given a finite sequence of nonnegative integers, consider the process of replacing each of the integers by the number of integers to its right which are smaller. Repeat the process in turn on each of the sequences obtained. Show that one will finally obtain a sequence such that repetition of the process makes no further change, i.e., each integer equals the number of smaller integers to its right.

E 1932. *Proposed by C. S. Ogilvy, Hamilton College*

For any point $P(x,y)$ on the curve $y=f(x)$, the slope of the line OP is equal to $f'(x_0)$, where $x_0=n^{1/(1-n)}x$. Find $f(x)$.

E 1933. *Proposed by H. Kestelman, University College, London, England*

Let $M_1, \cdots, M_k$ be $n\times n$ matrices whose sum is (δ_{rs}), and such that the sum of the ranks of the matrices is n. Prove that $M_rM_s=\delta_{rs}M_r(1\leqq r, s\leqq k)$.

E 1934. *Proposed by Gregory Dropkin and Brian Schmidt, Washington, D. C.*

Apply Newton's method for approximating the roots of a function [root$=\lim_{n\to\infty}x_n$, $x_{i+1}=x_i-f(x_i)/f'(x_i)$] to the polynomial $y=1+x^2$. What is the nature of the sequence $\{x_n\}$? For which choices, if any, of x_0 is the sequence bounded?

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before May 31, 1967.

5430. *Proposed by W. A. J. Luxemburg and R. F. Taylor, California Institute of Technology.*

Find all possible functions f: $R\to R$ with properties

(1) $f(x+y)=f(x)+f(y)$, and
(2) $f(p(x))=p(f(x))$ for some polynomial $p(x)$ of degree $\geqq 2$.

5431. *Proposed by George Schumm, Earlham College, Richmond, Indiana*

Given the ordinal numbers ξ and γ, we call ξ a root of γ if there exists an ordinal number $\alpha>1$ such that $\xi^\alpha=\gamma$. What is the kth smallest transfinite number for which there exists no root of degree greater than 1?

5432. *Proposed by Benjamin Volk, Yeshiva University*

Let $f(n)$ be an integer-valued function defined on the positive integers such that $f(n) \to 0$ (c, k). Show that

$$\sum_n \frac{1}{n} \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right) \to 0 \qquad (c, k).$$

5433. *Proposed by Benjamin Volk, Yeshiva University*

Let ϕ be a function, continuous on $(0, 1]$ of period 1 and satisfying

$$\sum_{j=1}^{n} \phi\left(x + \frac{j-1}{n}\right) = 0, \qquad n = 1, 2, \cdots.$$

Show that the sequence $\{\phi(n!x)\}$ is orthogonal on $(0, 1)$. (Cf. 5143 [1964, 1050]).

5434. *Proposed by H. Kestleman, University College, London, England*

Let N be a normal $n \times n$ matrix (i.e., N commutes with its hermitian transpose), $N^k = 1$ for some integer $k \geqq 2$, and $N \neq 1$. If $M = (1/k) \cdot (1 + N + N^2 + \cdots + N^{k-1})$ and $M \neq 0$, show that M projects orthogonally on to a subspace of complex-euclidean n-space, and identify this subspace in relation to N. (Note that if $n = k = 2$, then N is a reflection and $\frac{1}{2}(I + N)$ is the projection onto the mirror.)

5435. *Proposed by J. P. Williams, University of Michigan*

In the appendix to his book, *Fourier Analysis on Groups*, (Interscience, 1962), W. Rudin asserts that if A is a commutative Banach algebra then each nonzero complex homomorphism on A has norm 1. Is this true in general?

5436. *Proposed by Seymour Metz, Computer Associates, Arlington, Va.*

Does there exist a real function on the real line whose points of continuity form a dense set of measure zero?

5437. *Proposed by Hewitt Kenyon, George Washington University*

We say that a sequence x is eventually in A if $x_n \in A$ for all sufficiently large n; a sequence y is an undersequence to a sequence x if and only if whenever x is eventually in a set A, then y is eventually in A; we shall also say x is greater than y in this case. Two sequences will be called equivalent when each is an undersequence of the other.

Suppose that S is a nonempty set of sequences, every finite subset of which has a common undersequence. Which of the following statements are necessarily true?

(a) If y is an undersequence of x, then there exists a subsequence z of x which is equivalent to y.

(b) If S is finite, then S has a common subsequence.

(c) If S is finite, then S has a greatest common subsequence.

(d) If S is finite, then S has a greatest common undersequence.
(e) If S is countable, then S has a common subsequence.
(f) If S is countable, then S has a common undersequence.
(g) If S is countable, then S has a greatest common undersequence.
(h) S has a common undersequence.

5438. *Proposed by M. L. Glasser, Battelle Memorial Institute*

Express the following integrals in terms of Fresnel functions:

$$\int_0^\infty e^{x \sin y} \sin \left\{x \cos y \pm \tfrac{1}{2}(y + n\pi)\right\} dy.$$

5439. *Proposed by C. J. Mozzochi, University of Connecticut*

Let R be the reals with the usual topology, $\mathfrak{J}$, and $N(\Delta)$ the family of all neighborhoods of the diagonal Δ of $R\times R$ (i.e., $\Delta = \{(x, x): x\in R\}$). Show that $N(\Delta)$ is a uniformity for R, and $\mathfrak{J}$ is the uniform topology.

SOLUTIONS OF ADVANCED PROBLEMS

Integral Solutions of $3x^4-2y^2=1$

5255 [1965, 85]. *Proposed by J. H. E. Cohn, Bedford College, London, England*

Prove (or disprove): The equation in positive integers $2y^2=3x^4-1$ has solutions only if $x=1$ or $x=3$.

Editorial Note. No correct solutions were received for this problem. Simon Vatriquant provided a reference to Edouard Lucas, *Recherches sur l'analyse indéterminée et l'Arithmétique de Diophante*, 1961 reprint, Section VI, for rational solutions to the given equation. The fact that there are no integral solutions other than the ones cited is the subject of a report by R. T. Bumby. Bumby's abstract appearing in the Notices of the American Mathematical Society containing the program of the 71-st summer meeting (1966) follows:

J. H. E. Cohn has conjectured (see *Eight Diophantine Equations*, Proc. Lond. Math. Soc., XVI, pp. 153–166) that the equation $3x^4-2y^2=1$ has only the solutions $x=\pm 1$, $y=\pm 1$ and $x=\pm 3$, $y=\pm 11$. This conjecture is correct and may be demonstrated by extending Cohn's methods to the field generated by $(-2)^{1/2}$. Our solution requires the law of quadratic reprocity for this field.

Countable Subsets of E^n

5330 [1965, 1030]. *Proposed by M. Edelstein, Dalhousie University*

Let S be a subset of E^n and suppose that the distance $d(s_1, s_2)$ is rational for all pairs $s_1, s_2\in S$. Prove, or disprove, that S is countable.

Solution by P. R. Chernoff, Harvard University. The only relevant property of the rationals is their countability. By translating if necessary, we may assume $0=v_0\in S$. We may also assume that S lies in no proper subspace of E^n; hence, S contains a basis $v_1, v_2, \cdots, v_n$ of E^n. The intersection of $n+1$ spheres centered respectively at $v_0, v_1, \cdots, v_n$ can contain at most one point. Therefore, to a given $(n+1)$-tuple of admissible distances from $v_0, v_1, \cdots, v_n$ there can cor-

respond at most one point of S. Accordingly, S is countable.

Also solved by C. W. Anderson, Andreas Blass, W. G. Dotson, Jr., P. Erdös, R. L. Farrell, Bruce Feldmeyer, Leopold Flatto, G. J. Foschini, J. H. Foster, D. A. Gale, R. Goldstein (England), K. R. Goodearl, D. A. Hejhal, Ellen Hertz, E. C. Hook, David Huestis, Erwin Just, Irving Katz, D. C. Kay, T. J. Killeen, Betty Kvarda, W. M. Lambert, Jr., E. S. Langford, Sim Lasher, R. N Leggett, Jr., K. O. Leland, Bruce Lercher, Robert McGuigan, R. L Madell, J. G. Mauldon (England), M. D. Mavinkurve (India), J. C. Morgan II, Leo Moser & Mangesh Murdeshwar, Harsh Pittie, A. L. Port & Robert Cohen, George Purdy (England), Donald Quiring, S. M. Robinson, J. T. Rosenbaum, J. J. Schäffer (Uruguay), G. F. Schumm, A. L. Selman & Roger Hindley, D. L. Silverman, David Stanberry, Min Min Tang, N. R. Wallach, J. Ernest Wilkins, Jr., Margaret R. Wiscomb, and the proposer.

Erdös calls attention to Theorem III in his paper, *Some remarks on set theory*, Proc. Amer. Math. Soc., VI (1950), for a solution of the problem. He also recalls an unresolved question by Hajnal and himself: Let U be any countable set of real numbers. Join two points in E^n by an edge if their distance is in U. Is aleph-null the chromatic number of the resulting graph?

Iteration of Mappings in a Hilbert Space

5331 [1965, 1030]. *Proposed by M. Edelstein, Dalhousie University*

Let f be a mapping of l_2 into itself defined as follows. If

$$x=(x_1, x_2, \cdots, x_n, \cdots)\in l_2,$$

then $y=(y_1, y_2, \cdots, y_n, \cdots)=f(x)$ if, and only if, $y_n=e^{2\pi i/n!}(x_n-1)+1$. Is the sequence of iterates $f^n(0)$ of the origin bounded?

Solution by M. D. Mavinkurve, Siddharth College, India. By recursion the nth co-ordinate of the kth iterate of zero is obtained as $x_n^{(k)}=1-e^{2\pi ki/n!}$, so that the square of the norm of the kth iterate equals

$$4\sum_{n=1}^{\infty}\sin^2\frac{k\pi}{n!}=4F(k),$$

say, where $F(x)=\sum_{n=1}^{\infty}\sin^2(x\pi/n!)$. The series defining F converges uniformly on each finite interval and the series of derivatives converges uniformly on the whole line. Therefore F', which exists as the uniform limit of a sequence of periodic functions, is almost periodic; in particular F' is bounded. Therefore, if the sequence $\{F(k)\}$ were bounded, F itself would be bounded on the whole line (mean value theorem). Since

$$F(x)=\int_0^x F'(x)dx,$$

where F' is almost periodic, F also would be almost periodic and have a finite average value given by

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T F(x)=\lim_{T\to\infty}\sum_{r=1}^{\infty}\left(1-\frac{\sin(2\pi r/n!)}{2\pi r/n!}\right).$$

Such a limit, however, does not exist, for at $T=m!$ the value of the above sum exceeds m. Therefore the sequence of iterates is not bounded.

Also solved by D. A. Hejhal who obtains the unboundedness of $F(k)$ by direct calculation using k of the form $k=e_1 1!+e_2 2!+\cdots+e_N N!$, where $e_j=[(j+2)/3]$.

Fixed Point of a Set Operation

5332 [1965, 1030]. *Proposed by R. W. Forcade, Southern Illinois University*

Let T be an operator on (all) the subsets of a set X, such that the image of an arbitrary union of subsets is the intersection of their images. Prove that if every nonempty subset has a nonempty intersection with its image, then there is exactly one subset of X which is fixed by T.

Solution by Ellen Hertz, Columbia University. First we prove: If F operates on the subsets of X and $A\subseteq B$ implies $F(A)\subseteq F(B)$, then F has at least one fixed set. Let Z be the set of subsets of X such that $A\supseteq F(A)$. Z is not empty because $X\in Z$. Let R be the intersection of all sets of Z. If $A\in Z$ then $R\subseteq A$, so $F(R)\subseteq F(A)\subseteq A$. Then $F(R)\subseteq R$. But then $F(F(R))\subseteq F(R)$ so that $F(R)\in Z$. But R is contained in any member of Z, so $R\subseteq F(R)$. Therefore $R=F(R)$.

On the other hand, if $A\subseteq B$ implies $T(A)\supseteq T(B)$, then $T^2(A)\subseteq T^2(B)$, and we can apply the preceding result to T^2, letting A be the fixed set and $B=T(A)$. Then $B=T(A)$, $A=T(B)$. Write $A\triangle B=(A-B)\cup(B-A)$.

Now $T(A\cup B)=B\cap A$. Since $A\cap B\subseteq A$, $T(A\cap B)\supseteq B$. Similarly, $T(A\cap B)\supseteq A$. Then $T(A\cap B)\supseteq A\cup B\supseteq A\triangle B$. So we have $T(A\cup B)=T((A\cap B)\cup(A\triangle B))=T(A\cap B)\cap T(A\triangle B)$, and $A\cap B=T(A\cap B)\cap T(A\triangle B)$. It follows that $(A\cap B)\cap(A\triangle B)=T(A\cap B)\cap T(A\triangle B)\cap(A\triangle B)=T(A\triangle B)\cap(A\triangle B)$. But the left side is empty; therefore $A\triangle B$ is empty, i.e. $A=B$.

To establish uniqueness: Let $T(A)=A$, $T(B)=B$. Again we have $T(A\cup B)=A\cap B$, $T(A\cap B)\supseteq A\triangle B$, and the same reasoning shows that $A\triangle B$ is null.

Also solved by Roy O. Davies (England), Fred Galvin, Fred Galvin & J. C. Morgan II. J. Kahane & M. Tainiter, W. A. McWorter, J. G. Mauldon (England), M. D. Mavinkurve (India), A. Meir & J. S. W. Wong, J. C. Morgan II, C. B. A. Peck, Donald Quiring, J. T. Rosenbaum, P. A, Torelli & W. T. Smythe, and the proposer.

Several solvers found the fixed set as the union of sets A for which $A\subset T^2(A)$. Galvin and Morgan generalized to a lattice $\mathfrak{L}=\langle L,\cup,\cap\rangle$ (cf. G. Szasz, *Introduction to Lattice Theory*, p. 47) complete and containing a least element 0, and proved the following: If for every $x, y\in L$ for which $y\subseteq x$ there is an element $z\in L$ such that $y\cap z=0$, $y\cup z=x$, and T: $L\to L$ has the properties(1) $T(x\cup y)=T(x)\cap T(y)$, (2) $T(x)\cap x\neq 0$ if $x\neq 0$, then T has a unique fixed point.

An Application of the Arithmetic-Geometric Inequality

5333 [1965, 1030]. *Proposed by Masakazu Aoyagi, Chiba University, Japan*
If $0<x_1<x_2<\cdots<x_n\leq 1$, then prove that

$$\left(n\sum_{k=1}^{n} x_k\right)\Big/\left(\sum_{k=1}^{n} x_k+nx_1x_2\cdots x_n\right)\geq\sum_{k=1}^{n}\frac{1}{1+x_k}\cdot$$

I. *Solution by A. Meir, University of Alberta, Canada.* If $n=1$ the inequality

is false. Suppose $n \geqq 2$. By the inequality of means for any sequence $a_k > 0$, $\sum_{k=1}^n a_k \cdot \sum_{k=1}^n a_k^{-1} \geqq n^2$ and thus

$$\sum_{k=1}^n a_k \cdot \sum_{k=1}^n a_k^{-1}(1 - a_k) \geqq n \sum_{k=1}^n (1 - a_k).$$

Setting $a_k = x_k(1+x_k)^{-1}$ we obtain next

$$\sum_{k=1}^n x_k^{-1} \cdot \left(n - \sum_{k=1}^n (1 + x_k)^{-1} \right) \geqq n \sum_{k=1}^n (1 + x_k)^{-1}.$$

Multiplying both sides of this result by $x_1 x_2 \cdots x_n$ and observing that for $n \geqq 2$,

$$\sum_{k=1}^n x_k \geqq x_1 x_2 \cdots x_n \cdot \sum_{k=1}^n x_k^{-1};$$

we obtain the proposed inequality, after rearranging terms.

II. *Solution by David Borwein, University of Western Ontario.* As the result is evidently false when $n = 1$, suppose that $n \geqq 2$. Let

$$A_r = \frac{1}{r} \sum_{k=1}^r x_k,\ H_r = \frac{1}{r} \sum_{k=1}^r \frac{1}{1 + x_k},\ G_r = (x_1 x_2 \cdots x_r)^{1/r},\ G_0 = 1;$$

and suppose that $x_1, x_2, \cdots, x_n$ satisfy

$$(1) \qquad 0 < x_1 \leqq x_2 \leqq \cdots \leqq x_n \leqq \frac{1}{G_{n-1}}$$

instead of the more stringent condition in the original problem. It is easily deduced from (1) that

$$(2) \qquad G_r \leqq 1 \quad \text{and} \quad x_r G_{r-1} \leqq 1 \qquad (r = 1, 2, \cdots, n).$$

We shall prove the proposition

$$P(r): \qquad H_r \leqq \frac{1}{1 + G_r}$$

to be true for $r = n$. Since $A_n \geqq G_n$, it follows from $P(n)$ and the first part of (2) that

$$\frac{A_n}{A_n + G_n^n} \geqq \frac{1}{1 + G_n^{n-1}} \geqq H_n,$$

which yields the required inequality.

We first prove the

LEMMA: *If* $0 < a \leqq b$, $a + b = 1$, $0 < g \leqq 1$ *and* $0 < x - g \leqq 1$, *then*

$$\frac{a}{1+x}+\frac{b}{1+g}\leqq\frac{1}{1+x^a g^b}\,.$$

Differentiating

$$f(x)=\frac{1}{1+x^a g^b}-\frac{a}{1+x}-\frac{b}{1+g}$$

and simplifying, we find that $f'(x)$ has the same sign as

$$(x^{(1+a)/2}g^{b/2}-1)((g/x)^{b/2}-1),$$

so that $f'(x)\geqq 0$ when $g\leqq x\leqq g^{-b/(1+a)}$ and $f'(x)\leqq 0$ otherwise. Hence $f(x)\geqq f(g)=0$ when $0<x\leqq g^{-b/(1+a)}$; and, when $g^{-b/(1+a)}\leqq x\leqq g^{-1}$,

$$f(x)\geqq f(g^{-1})=\frac{1}{1+g^{b-a}}-\frac{ag+b}{1+g}=\frac{\phi(g)}{(1+g^{b-a})(1+g)},$$

where $\phi(g)=a+bg-ag^{1+b-a}-bg^{b-a}$.

To complete the proof of the lemma it suffices to show that $\phi(g)\geqq 0$. Differentiating $\phi(g)$ twice in the range $0<g\leqq 1$, we see that $\phi''(g)\geqq 0$ so that $\phi'(g)\leqq\phi'(1)=0$ and hence $\phi(g)\geqq\phi(1)=0$.

We can now show that $P(n)$ is true by induction. $P(1)$ is trivially true: if $P(r-1)$ is true for any r in the range $\{2,\cdots,n\}$ then, by (2) and the lemma with $x=x_r$, $g=G_{r-1}$, $a=1/r$, $b=1-1/r$, there follows

$$\begin{aligned}H_r&=\frac{1}{r(1+x_r)}+\left(1-\frac{1}{r}\right)H_{r-1}\\&\leqq\frac{1/r}{1+x_r}+\frac{1-1/r}{1+G_{r-1}}\\&\leqq\frac{1}{1+x_r^{1/r}G_{r-1}^{1-1/r}}=\frac{1}{1+G_r},\end{aligned}$$

so that $P(r)$ is true.

Also solved by Max A. Bershad, N. P. Bhatia, P. H. Diananda (Singapore), R. F. Farrell, M. D. Mavinkurve (India), Edward Saibel, P. S. Schnare, and the proposer.

The Determinant of a Real Symmetric Matrix

5334 [1965, 1030]. *Proposed by L. A. Shepp, Bell Telephone Laboratories, Murray Hill, N. J.*

If $(b)=[b_{ij}]$ is a real symmetric square matrix with $b_{ii}=1$, and $\sum_{j\neq i}|b_{ij}|\leqq 1$ for each i, then $\det(b)\leqq 1$.

Solution by Stanton Philipp, University of California, Riverside. Let $\lambda_1,\cdots,\lambda_n$ be the characteristic roots of (b). Suppose that λ is a characteristic root of

(b) and let $x=(x_1, \cdots, x_n)^T$ be an associated (nonzero) characteristic vector. Let x_k be one of the x_i which is largest in absolute value. Then

$$\sum_j b_{kj}x_j = \lambda x_k, \quad (\lambda - 1)x_k = \sum_{j\neq k} b_{kj}x_j, \quad \text{and} \quad |\lambda - 1| \leqq \sum_{j\neq k} |b_{kj}| \leqq 1.$$

Since (b) is real and symmetric, λ is real and $0 \leqq \lambda$. Notice that trace $(b)=n = \sum_i \lambda_i$. Employing the arithmetic-geometric mean inequality we have

$$\det(b) = \lambda_1\lambda_2 \cdots \lambda_n \leqq \left(\frac{1}{n}\sum_i \lambda_i\right)^n = 1.$$

Also solved by D. E. Crabtree, W. R. Gordon, J. Z. Hearon, Irving Katz, Marvin Marcus, M. D. Mavinkurve (India), R. K. Meany, A. Meir & J. S. W. Wong, P. L. Nikolai, J. T. Rosenbaum, Edward Saibel, A. J. Schneider, V. Seshadri, Richard Sinkhorn, Sidney Spital, J. H. van Lint, S. Zlobec (Yugoslavia), and the proposer.

Several solvers note the validity of the conclusion for Hermitian matrices. The inequality $|\lambda-1| \leqq 1$ is a special case of Gerschgorin's theorem (M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, 1964, p. 146).

A Closed Set of Idempotents

5335 [1965, 1030]. *Proposed by James Chew, Virginia Polytechnic Institute*

In this MONTHLY (Dec. 1963) Doyle and Warne, *Some properties of groupoids*, the following theorem is given without proof: *The set of idempotents in a topological groupoid is* (*topologically*) *closed.* Supply a proof.

Solution by W. G. Dotson, Jr., North Carolina State University, Raleigh. This result follows immediately from the definition (as given by Doyle and Warne) of a topological groupoid as a groupoid which is a Hausdorff topological space in which the group multiplication is continuous in both variables. For, if the set E of idempotents in a topological groupoid G is not topologically closed, then there is a limit point x of E such that $x^2 \neq x$. There exist disjoint open sets $V(x^2)$, $V(x)$ containing x^2, x respectively. Since the multiplication is continuous, there exist open sets $S(x)$, $T(x)$, each containing x, such that $S(x)T(x) \subset V(x^2)$. But $S(x) \cap T(x) \cap V(x)$ is an open set containing x and so contains an element $e \in E$. Hence $e^2 \in V(x^2)$ whereas $e \in V(x)$, so that $e^2 \neq e$, a contradiction.

Also solved by Shair Ahmad, P. R. Chernoff, C. C. Clever, A. J. D'Aristotle, L. L. English, D. A. Hejhal, Harsh Pittie, K. N. Sigmon, J. F. Standish, Kenneth Yanosko, and the proposer.

Pittie generalizes by proving that the set of fixed points of a continuous map f: $X \to X$ (X Hausdorff) is closed.

A Sum over k-tuples

5336 [1965, 1030]. *Proposed by D. J. Newman, Yeshiva University*

Let k be an integer greater than 1. Prove that

$$\sum \frac{1}{n_1 n_2 \cdots n_k(n_1 + n_2 + \cdots + n_k)} = k!,$$

the sum extending over all k-tuples of positive integers which have no nontrivial common factor.

Solution by J. H. van Lint, Bell Telephone Laboratories, Murray Hill, N. J. Let S be the sum to be determined. Then

$$\sum_{n_i \geq 1} \frac{1}{n_1 n_2 \cdots n_k(n_1 + n_2 + \cdots + n_k)} = S \sum_{d=1}^{\infty} \frac{1}{d^{k+1}} = S\zeta(k+1).$$

The sum on the left-hand side is equal to

$$\sum_{\nu=k}^{\infty} \sum_{n_1+\cdots+n_k=\nu} \frac{1}{n_1 n_2 \cdots n_k(n_1 + \cdots + n_k)} = \int_0^1 \left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots\right)^k \frac{dx}{x}$$

$$= \int_0^1 \left(\log \frac{1}{1-x}\right)^k \frac{dx}{x} = \int_0^{\infty} \frac{u^k}{e^u - 1} du = \int_0^{\infty} u^k \left(\sum_{n=1}^{\infty} e^{-nu}\right) du$$

$$= \sum_{n=1}^{\infty} \frac{\Gamma(k+1)}{n^{k+1}} = k!\zeta(k+1).$$

Hence $S = k!$.

Also solved by L. Carlitz, N. J. Fine, J. H. Halton, E. C. Milner & A. Oppenheim (England), and the proposer.

Carlitz observes that a similar proof yields

$$\sum_{n_i \geq 1} \frac{1}{n_1 n_2 \cdots n_k(n_1 + \cdots + n_k + 1)} = k!,$$

the sum being extended over all positive integral k-tuples.

Ratio of Sums of Euler's Totient Function

5337 [1965, 1031]. *Proposed by C. S. Venkataraman, Sri Kerala Varma College, Trichur, India*

Evaluate $\sum \phi(n/e)/\sum \phi(n/d)$, where e runs through the even divisors and d runs through the odd divisors of n, and ϕ is Euler's function.

Solution by L. E. Clarke, University of East Anglia, Norwich, England. We show more generally that, for any positive integer m,

$$S_n = \sum{}' \phi(n/d)/\sum{}'' \phi(n/d) = \{g - \phi(g)\}/\phi(g),$$

where $\sum'(\sum'')$ denote summation over those divisors d of n for which $(d, m) > 1$ ($(d, m) = 1$), and $g = (m, n)$. On putting $m = 2$, it follows that the expression given in the problem is 0 or 1 according as n is odd or even.

Suppose n is expressed in the form $\prod p_i^{a_i}$, and write

$$n_1 = \prod_{p_i \nmid m} p_i^{a_i}, \qquad n_2 = \prod_{p_i \mid m} p_i^{a_i}.$$

Then

$$\sum{}'' \phi(n/d) = \sum_{d|n_1} \phi(n_1 n_2/d) = \phi(n_2) \sum_{d|n_1} \phi(n_1/d) = n_1\phi(n_2)$$

and

$$S_n = \left(n - \sum{}''\right) \Big/ \sum{}'' = \{n_2 - \phi(n_2)\}/\phi(n_2) = \{g - \phi(g)\}/\phi(g)$$

because g and n_2 have the same prime factors.

Also solved by P. N. Bajaj, M. A. Bershad, L. Carlitz, S. R. Cavior, N. J. Fine, Sylvan Greene, M. G. Greening (Australia), Emil Grosswald, E. M. Horadam (Australia), F. T. Howard, Bernard Jacobson, M. S. Kaplan, Irving Katz, D. A. Klarner, Marijo Levan, C. C. Lindner, H. London, Ka Menehune, Lieselotte Miller, F. R. Olson & Al Somayajulu, J. C. Parnami (India), C. B. A. Peck, Stanton Philipp, Harsh Pittie, R. W. Prielipp, D. A. Prener, Simeon Reich (Israel), Margaret B. Seay, Sister Marion Beiter, R. Sivaramakrishnan (India), Sidney Spital, E. W. Trost (Switzerland), A. M. Vaidya (India), J. H. van Lint, C. S. Venkataraman (India), and Jean Williams.

Sivaramakrishnan proves the generalization: If $f(n)$ is any multiplicative arithmetic function and $F(n)=\sum_{d|n} f(d)\neq 0$, then $\sum f(n/e)/\sum f(n/d)=0$ if n is odd, and $=F(2^{\nu-1})/f(2^\nu)$ where ν is the exponent of 2 in n if n is even.

Continuity of a Mapping from a Disk to its Boundary

5338 [1965, 1031]. *Proposed by J. W. Wyman, Pasadena College*

Let f: $D\rightarrow D$ be a continuous function on the unit disk D in the complex plane such that $f(z)\neq z$ for each $z\in D$. Let ∂D denote the boundary of D and let ψ: $D\rightarrow D$ be defined as follows: If $z\in D$, the line determined by $f(z)$ and z intersects ∂D in exactly two places. Let $\psi(z)$ be the point on ∂D such that $f(z)$, z and $\psi(z)$ are in that order. Then ψ is continuous. Supply a proof. (This result is stated without proof in Tucker, *Some topological properties of disk and sphere*, Proceedings of the First Canadian Mathematical Congress, 1946, pp. 285–309.)

I. *Solution by Dennis A. Hejhal, student, Lane Technical High School, Chicago, Ill.* Choose any $\zeta\in D$ and draw the ray from $f(\zeta)$ to ζ to $\psi(\zeta)$ on $|z|=1$. On $|z|=1$ draw arcs of measure ϵ in both directions from $\psi(\zeta)$ and let P_1, P_2 be the endpoints. From P_1, P_2 draw segments through R, the midpoint of ζ, $f(\zeta)$, to Q_1, Q_2 respectively on ∂D. Consider the circle $|z-f(\zeta)|=\eta$ entirely within D and angle Q_1RQ_2. Choose δ so that $|z-\zeta|<\delta$ is within D and angle P_1RP_2, and so that $|f(z)-f(\zeta)|<\eta$ for all z, $|z-\zeta|<\delta$. It follows therefore that if $|z-\zeta|<\delta$, the ray joining $f(z)$ to z cuts ∂D at a point between P_1 and P_2, implying the continuity of $\psi(z)$.

II. *Solution by C. R. Deeter, Texas Christian University.* The line determined by z and $f(z)$ for fixed z is given by

$$l(t) = f(z)(1-t) + zt, \qquad t \text{ real}.$$

Then $\psi(z)=l[T(z)]$, where $T(z)$ is the positive solution of the quadratic equa-

tion $|l(t)|^2-1=0$. The coefficients of this quadratic are all continuous functions of z, and the leading coefficient, $|f(z)-z|^2$, cannot vanish since $f(z)\neq z$ for each $z\in D$. Since, in this case, a solution of the quadratic equation is a continuous function of its coefficients, $T(z)$ is a continuous function of z, as must also be $\psi(z)$.

Also solved by P. R. Chernoff, W. G. Dotson, Jr., J. H. Foster, K. R. Goodearl, P. S. Schnare, J. H. van Lint, Kenneth Yanosko, and the proposer.

Extremals for a Product of Distances

5339 [1965, 1031]. *Proposed by Necdet Ücoluk, Purdue University*

Let ABC be a given triangle in the plane, with a nonempty interior. Determine the position of a point N of the plane for which the product $NA\times NB\times NC$ attains an extreme value.

Solution by Michael Goldberg, Washington, D. C. Expressed analytically, the product surface is a sixth degree equation in the variables x, y, z. However, it is not necessary to resort to analysis to find the extreme values. For remote points, the product tends to infinity. At the vertices A, B, C the products are zero; for neighboring points, the products are positive. Hence, extreme values are attained at A, B, and C.

On the interior of the triangle, there is a minimax point at which the tangent to the product surface has zero slope, but the point is not an extreme point.

The product surface can be described as an udder with three teats. Each teat touches a vertex at the plane $z=0$. Below the udder, a section $z=c$ cuts the surface in three disconnected closed curves. These three curves are joined at a horizontal tangent at the udder. Above this tangent plane, a horizontal section is a single closed curve with three lobes.

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: KENNETH O. MAY, University of Toronto and E. P. VANCE, Oberlin College

Materials intended for review should be sent to Kenneth O. May, American Mathematical Monthly, Department of Mathematics, University of Toronto, Toronto 5, Ont., Canada.

Introduction to Lattice Theory. By D. E. Rutherford. Hafner, New York, 1965. x+117 pp. $5.75.

This little book is intended to introduce the undergraduate to the simpler parts of lattice theory and with the exception of some of the applications of Boolean algebras all of the topics covered appear as easier parts of the first twelve chapters of G. Birkhoff, *Lattice Theory*, Amer. Math. Soc. Colloq. Publ., Vol. 25, rev. ed. Thus the definitions and elementary properties of modular,

metric, and distributive lattices, Boolean algebras, semi-modular and matroid lattices are discussed in a leisurely and readable fashion; simple illustrations are provided and the classical applications to algebra, point set topology, geometry, Boolean and Brouwerian logic, and switching circuit theory are presented. Fifty-one fairly well-chosen exercises are included at the end of the text.

A large part of the book is devoted to the above mentioned applications of Boolean algebra, especially to switching circuits and it is here that the author is at his best. In particular the application of Boolean matrices and determinants to electrical networks is doubly attractive because of the rarity of the material in elementary texts.

On the other hand the book has shortcomings both in content and in form which prevent it from being considered a really well-balanced introduction to the subject. Of the former perhaps the most remarkable is the total absence of transfinite methods; thus, for example, the representation theorems for distributive lattices and Boolean algebras are discussed for the finite case only. Of the latter the poor discussion of free lattices is one example. Indeed, the definition of a free lattice given here is simply incorrect. Another example is the failure to distinguish between sublattices of direct products and subdirect products. There are others.

It seems to the reviewer that, at least for better undergraduates, judicious guided reading of Birkhoff, perhaps with this book as an aid, is still the best introduction to lattice theory.

A. F. PIXLEY, Harvey Mudd College

Special Functions and their Applications. By N. N. Lebedev. Translated and adapted from the Russian by R. A. Silverman. Prentice-Hall, Englewood Cliffs, N. J., 1965. xii+308 pp. $12.00.

The special functions here are the gamma function, the exponential, probability, logarithmic and Fresnel integrals, the classical orthogonal polynomials, spherical harmonics, cylinder and parabolic cylinder functions, and the Gaussian and confluent hypergeometric functions. The selection of functions and properties was evidently made with an eye to classical physics. The properties are developed with a careful use of complex variable techniques and, for those techniques that are not explained in detail, the reader is given specific references.

The applications (one-fifth of the book) are mostly boundary value problems. Applications of such complexity require of the reader some background in mathematical physics, but a large section of the material is a self contained unit of different Dirichlet problems.

Covering similar materials, this book is fuller and longer than Hochstadt's seventy five page book (this MONTHLY, 69 (1962) 325), and more motivated by physical applications than Rainville's book (this MONTHLY, 67 (1960) 1044).

Professor Silverman's translation reads well and he has furnished references easily available in English. Problem sets are included.

DAVID DICKINSON, University of Massachusetts

The Elements of Computational Mathematics. Edited by S. B. Norkin, English translation by A. D. Booth. Macmillan, New York, 1965. 192 pp. $6.00.

This small but relatively expensive text contains the following table of contents: Chapter I, Computation with approximate numbers, and assessment of errors; Chapter II, Construction of tables of functions; Chapter III, Approximate solution of equations; Chapter IV, Systems of linear equations; Chapter V, Interpolation polynomials; Chapter VI, Approximate computation of integrals; Chapter VII, Approximate integration of differential equations.

Except for Chapter VII the material could be presented to an Advanced Placement High School Class by a teacher well trained in the calculus. The notion of a derivative, Rolle's theorem and the mean value theorem for derivatives are used only occasionally. Definite integrals are treated by approximating, numerically, the limit of a sum.

Emphasis is placed on the understanding and use of algorithms. Unfortunately, practise is advocated by hand or by desk calculator and little account is taken of modern digital computing machinery. This is unfortunate, because the "elements of computational mathematics" are dull as dishwater. A fascination for them can be aroused by developing them in the environment of modern computing machinery.

The material in this text is developed with great care and is supported by good problems.

R. H. OWENS, University of Virginia

Problèmes de Mathématiques. By Daniel Dumas de Rauly. Gauthier-Villars, Paris, 1963. 221 pp. N.F. 16.

Here are 213 pages of solved problems, classified into 14 chapters. Compared to Schaum's Outlines, the problems are nonroutine, and none is merely repetitive. No problem in the book would seem out of place in an American "Advanced Mathematics for Engineers," and such a book should include many problems of this sort as examples or exercises. Differential equations are omitted on the grounds that there is an ample supply of problems on this topic.

The first two chapters present good problems from elementary combinatorial analysis and probability. The third gives a few superficialities from "modern algebra," apparently to back up a claim in the preface. The next two present standard topics from elementary linear algebra. The remaining two-thirds of the book is concerned with classical calculus of functions of from one to three variables. Perhaps the most ambitious part is two chapters devoted to the complex Gamma and Beta functions, with some resort to: "It can be shown that . . .".

Usually the author states a problem, tacitly assumes the relevant definitions and theorems to be known to the reader, and outlines a fairly complete solution, omitting details that should be clear. There are a good many misprints.

BURROWES HUNT, Reed College

Introduction to Abstract Algebra. By Roy Dubisch. Wiley, New York, 1965. ix+193 pp. $5.95.

This very elementary, but beautifully written text, constructs the usual number systems from the natural numbers, and then uses their properties to motivate the definitions of abstract algebraic structures. First, a list of eight axioms for the natural numbers is presented. (The verification that these axioms are equivalent to Peano's axioms is left for the student as a starred exercise.) The integers are constructed, their basic properties proved, and then there follows a chapter on integral domains and rings. Factorization in integral domains is discussed, and it is shown that unique factorization fails in a subdomain of the integers in $Q(\sqrt{5})$. Next, the rationals are constructed, followed by a chapter on fields and groups. As further examples of fields, the reals and complexes are constructed, and their characteristic properties clearly and thoroughly discussed, although the completeness of the reals is never proved. The quaternions are mentioned also, as an example of a skew-field.

There is a short chapter on vector spaces and a very nice one on polynomial rings. After proving unique factorization for polynomials with coefficients in a field, the author gives the solutions to the general equations of the third and fourth degree, and then discusses the impossibility of finding solutions to the general equations of higher degrees.

The author's teaching talent is evident throughout the text. By the use of many carefully chosen examples, he conscientiously warns his readers of many common misunderstandings. The book contains many exercises, most of which are either prescribed computations to illustrate abstract ideas, or completions of easy proofs not entirely done in the text. Each chapter concludes with a few "thought questions," many of which seem too difficult for students at this level, but each is accompanied by an explicit reference to the complete bibliography of 63 titles.

One might hope for more theorems within the discipline of abstract algebra itself. The most abstract proof given in the text is the verification that the automorphisms of a group (or a field) themselves form a group.

WELLS JOHNSON, Bowdoin College

Asymptotic Expansions. By E. T. Copson. Cambridge Tracts in Mathematics and Mathematical Physics, No. 55, Camb. Univ. Press, New York, 1965. vii+120 pp. $6.00.

This monograph gives an introduction to the theory of asymptotic representation of a function defined by a definite integral or a contour integral. A knowledge of the elementary parts of the theory of functions is assumed, in particular complex integration and special functions such as the Gamma function and Bessel functions. After a preliminary account of the properties of asymptotic series, the standard methods of deriving the asymptotic expansion of an integral are explained in detail.

The chapter headings are: (1) Introduction (Historical remarks), (2) Pre-

liminaries, (3) Integration by parts, (4) The method of stationary phase, (5) The method of Laplace, (6) Watson's lemma, (7) The method of steepest descents, (8) The saddle-point method, (9) Airy's integral, (10) Uniform asymptotic expansions.

Chapters 5–7 contain applications to the Gamma function, the logarithmic integral, Bessel functions and the error function. There are no exercises and the applications of asymptotic methods to differential equations are not considered. As mentioned by the author in the preface, there are few theorems and the aim is to explain the available methods and to illustrate them by means of a few of the more important special functions. The typography is splendid. The reviewer feels that this monograph is a welcome addition to the existing literature on the subject.

S. M. SHAH, University of Kansas

Distribution Theory and Transform Analysis. By A. H. Zemanian. McGraw-Hill, New York, 1965. xiii+371 pp. $13.75.

Since a knowledge of distribution theory is now almost a necessary prerequisite for study in the systems area, and many other areas of applied mathematics, it is gratifying to see the appearance of this relatively elementary text directed to applied scientists.

The text is oriented toward the use of distributions in transform analysis and stress is generally placed on important concepts relevant to this goal. The presentation should generally be excellent for mathematically oriented science and engineering first year graduate students. The use of stars and diamonds to indicate important sections and results should facilitate reader use. Particularly commendable are the emphasis on L. Schwartz's theory, the stress on convolution, the introduction of ultradistributions to obtain the Fourier transform of all distributions, an application to the characterization of passive systems, and a treatment of periodic distributions. Considering the author's intent the shortcomings of the book are relatively minor. The most important of these seems the absence of the kernel theorem [I. M. Gel'fand, N. Vilenkin, "Generalized Functions," vol. 4, Academic Press, 1964, p. 18], which is of fundamental importance in linear systems theory. The fact that Mikusiński's calculus, p. 169, is a special application of a general result on embedding a ring in a field is not mentioned. Although Schwartz's notation is generally used, it is sometimes mixed up, as illustrated by $\mathfrak{D}_+'$ (which denotes support in $0 \leqq t < \infty$ in Zemanian but support bounded on the left in Schwartz). The results of Chapter 10 must be interpreted with care since a fixed physical system will have different properties depending upon the mathematical domain of definition chosen (for example, a parallel R-C circuit allows $i=f=0$, $v=\exp[-t/RC] \in \mathfrak{D}'$ in contrast to Lemma 3).

In summary the text is excellent and well-written for its purpose. Although most engineering instructors will have to supply more physical concepts, this should be an enjoyable and profitable text from which to teach.

R. W. NEWCOMB, Stanford University

Introduction to Field Theory. By Iain T. Adamson. Oliver & Boyd, Edinburgh and London, Interscience, New York, 1964. viii+180 pp. $2.75.

This is an attractive book on field theory and Galois theory at a level somewhat beyond that of Artin's *Galois Theory*. Some knowledge of group theory is assumed and the book is addressed to honors and graduate students. The fundamental theorem of finitely generated abelian groups and some elementary facts about prime power order groups are freely used.

There are four chapters: Chapter 1 introduces rings, fields and vector spaces as well as developing the basic facts about them; Chapter 2 is devoted to extensions of fields, with particular care given to field identifications; the basic theorems concerning the division algorithm and the existence of algebraic closures are stated without proofs; Chapter 3 is devoted to the Galois theory of finite dimensional Galois extensions and mirrors Artin's approach; Chapter 4 offers applications of the previous material and includes material on cyclotomic and cyclic extensions, Wedderburn's theorem on finite division rings, an especially complete discussion of constructibility by compass and straight-edge, and the impossibility of solving by radicals the general equation of degree >4.

The text is very clearly written, with many examples, and the exercises are good though there are too few for my taste. All told this is an excellent introduction to field theory.

NEIL GRABOIS, Williams College

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

University of California, San Diego: Associate Professor Adriano Garsia, California Institute of Technology, has been appointed Professor; Dr. Jay Fillmore, University of Minnesota, has been appointed Assistant Professor.

DePaul University: Dr. J. I. Goldman, Illinois Institute of Technology, has been appointed Assistant Professor; Assistant Professor Alphonse Buccino will be on leave for 1966–67 as an NSF Science Faculty Fellow at the University of Chicago.

Michigan State University: Dr. L. G. Woodby, U. S. Office of Education, Washington, D. C., has been appointed Professor; Associate Professor W. M. Fitzgerald, Florida Atlantic University, and Dr. C. J. Martin, Avco Corporation, Wilmington, Massachusetts, have been appointed Associate Professors; Assistant Professor G. D. Taylor, University of Arizona, and Drs. G. D. Ludden, University of Indiana, C. R. MacCluer, University of Michigan, C. E. Weil, University of Chicago, have been appointed Assistant Profes-

sors; Dr. I. E. Vance, University of Michigan, has been appointed Visiting Assistant Professor; Associate Professor John Wagner has been promoted to Professor.

Ohio Wesleyan University: Professor R. L. Wilson has been appointed Visiting Professor of Mathematics and Director of the Computing Center at the University of Ibadan, Nigeria, for the period 1966–68; Associate Professor R. V. Mendenhall has been promoted to Professor and appointed Chairman of the Mathematics Department; Professor S. E. Ganis has been appointed Director of the Computing Center.

University of South Carolina: Associate Professor W. L. Allen, Lamar State College of Technology, has been appointed Associate Professor; Mr. H. E. Scheiblich, University of Texas, has been appointed Assistant Professor.

Washington State University: Assistant Professors R. A. Stoltenberg and J. S. Rue, University of Wyoming, and Drs. D. C. Barnes, University of California, Davis, J. E. Cude, University of Texas, R. L. Irwin, University of Marburg, Germany, C. B. Millham, Iowa State University, have been appointed Assistant Professors; Dr. Elizabeth M. Strohmeier has been promoted to Assistant Professor; Assistant Professors W. M. Cunnea and J. H. Jordan have been promoted to Associate Professors.

Associate Professor L. R. Bragg, Case Institute of Technology, has been appointed Professor at Oakland University.

Mr. H. L. Carlson, University of Wisconsin, Madison, has accepted an appointment as Assistant Computer Scientist in the Applied Mathematics Division of Argonne National Laboratory.

Assistant Professor B. J. Cerimele, Xavier University, will be on leave during 1966–67 and has accepted a post-doctoral fellowship for advanced study and research in the field of biomathematics at the North Carolina State University at Raleigh.

Dr. C. A. Hall, White Sands Missile Range, New Mexico, has been appointed Senior Mathematician for the Bettis Atomic Power Laboratory of Westinghouse Electric Corporation.

Dr. H. B. Keynes, Wesleyan University, has been appointed Assistant Professor at the University of California at Santa Barbara.

Professor W. M. Perel, University of North Carolina, has been appointed Professor at Wichita State University.

Dr. A. K. Ray, Clarkson College of Technology, has been appointed Associate Professor at the University of Ottawa.

Professor Alex Rosenberg, Cornell University, has been appointed Chairman of the Mathematics Department.

Associate Professor J. L. Smith, Muskingum College, has been appointed Chairman of the Mathematics Department.

Dr. H. P. Thielman, Teledyne, Inc., Alexandria, Virginia, died on February 17, 1966. He was a member of the Association for 30 years.

Professor J. R. Vatnsdal, Washington State University, died on April 14, 1966. He was a member of the Association for 20 years.

FELLOWSHIP AND RESEARCH OPPORTUNITIES IN MATHEMATICS

The Division of Mathematical Sciences, National Academy of Sciences—National Research Council, calls attention to a variety of fellowship and other support for research in the mathematical sciences at both the predoctoral and postdoctoral levels to be awarded during the year 1966–67. Copies of the complete announcement are available from the Division of Mathematical Sciences, National Academy of Sciences—National Research Council, 2101 Constitution Avenue, Washington, D. C. 20418.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

THE FORTY-SEVENTH SUMMER MEETING OF THE ASSOCIATION

The Forty-seventh Summer Meeting of the Mathematical Association of America was held at Rutgers—The State University, New Brunswick, New Jersey, from Monday, August 29, through Thursday, September 1, 1966, in conjunction with summer meetings of the American Mathematical Society, the Society for Industrial and Applied Mathematics, the Institute of Mathematical Statistics, the Pi Mu Epsilon Fraternity and Mu Alpha Theta. There were registered 1856 persons, including 926 members of the Association.

Sessions of the Association were held on Monday morning and afternoon, on Tuesday morning, on Wednesday afternoon, and on Thursday afternoon. All sessions were held in Room 123 of Scott Hall at Rutgers; the proceedings could also be viewed in Room 135 of Scott Hall by means of closed circuit television on a large screen. Presiding officers at the three Earle Raymond Hedrick Lectures were President R. L. Wilder, First Vice-President G. S. Young, and Second Vice-President A. B. Willcox; at the two lectures on Monday morning, Professor H. F. Trotter; at the session on Monday afternoon, Professor R. D. Luce; at the panel discussion on Tuesday morning, Professor Joshua Barlaz; at the lecture on Wednesday afternoon, Professor Mary P. Dolciani; and at the session on Thursday afternoon, Professor Everett Pitcher. The fifteenth series of Earle Raymond Hedrick Lectures was delivered by Professor N. J. Fine of Pennsylvania State University. The Program Committee consisted of Joshua Barlaz, Chairman; Mary P. Dolciani, R. D. Luce, A. E. Meder, Jr., Albert Nijenhuis, Everett Pitcher, and H. F. Trotter.

FIRST SESSION OF THE ASSOCIATION

The Earle Raymond Hedrick Lectures: *Basic Hypergeometric Series and Applications*, Lecture I, by Professor N. J. Fine, Pennsylvania State University.

The theory of partitions, founded by Euler, leads naturally to series of the form

$$F(a, b; t:x) = 1 + \frac{(1-ax)}{(1-bx)}t + \frac{(1-ax)(1-ax^2)}{(1-bx)(1-bx^2)}t^2 + \cdots \tag{*}$$

There are simple relations connecting (*) with each of $F(ax, b; t:x)$, $F(a, bx; t:x)$, and $F(a, b; tx:x)$. From them one can obtain famous identities of Euler, Gauss, and Jacobi, as well as other new ones. Suitable generalizations, specializations, and paraphrases yield results on partitions, sums of squares, Liouville identities, Gaussian sums, elliptic functions, modular equations, and mock-theta functions. One example of a new result is the following. Let $f(x) = (1-x)(1-x^2)(1-x^3)\cdots$. Then

$$\frac{f(x^2)f(x^3)f(x^8)f(x^{12})}{f(x)f(x^{24})} = 1 + A_1x + A_2x^2 + \cdots,$$

where A_n is the excess of the number of divisors of n which are congruent to 1, 5, 7, 11 (mod 24) over the number congruent to -1, -5, -7, -11. Applications have also been made to orthogonal functions, enumeration problems in finite fields, statistics, and even mathematical psychology.

Convergence of Fourier Series, by Professor V. F. Cowling, Rutgers—The State University.

In his thesis "The integral and trigonometric series," written in 1915, N. N. Lusin expressed the conjecture that the Fourier series of functions in L^2 converge almost everywhere. Recently L. Carleson (in a manuscript entitled "On the growth of partial sums of Fourier series" unpublished as of the present writing) disposed of this conjecture by means of the following theorem:

(a) If for some $\delta>0$ $\int_{-\pi}^{\pi} |f(x)| (\log^+|f(x)|)^{1+\delta}dx<\infty$, then $S_n(x) = o(\log\log n)$, a.e.

(b) If $f\in L^p$, $1<p<2$, then $S_n(x)=o$ (log log log n), a.e. (c) If $f\in L^2$, then $S_n(x)$ converges a.e. In the case of part (c) which provides an affirmative answer to the Lusin conjecture the best previous result is the Kolmogorov-Seliverstov-Plessner theorem to the effect that $S_n(x)=o\ (\log n)^{1/2}$, a.e. One may demonstrate by means of Fejér's polynomials the existence of continuous functions in which the Fourier series converges at an everywhere dense set or in a set having the power of the continuum. Recently J. P. Kahane and Y. Katznelson (in a presently unpublished manuscript entitled "Sur les ensembles de divergence des séries trigonométriques") have shown that given a set of zero measure there exists a continuous function whose Fourier series diverges on this set.

The Continuum Problem, by Professor R. M. Smullyan, Yeshiva University.

This was an expository account of the work of Kurt Gödel (1938) and Paul Cohen (1963) concerning the independence of C.H. (the Generalized Continuum Hypothesis) relative to the axioms of Z.F. (Zermelo-Fraenkel Set Theory). The Continuum Hypothesis asserts that for any infinite cardinal C, $2^C=C^+$, where C^+ is the first cardinal larger than C. To this day, the truth or falsity of C.H. is unknown. The consistency of C.H. relative to Z.F. was demonstrated by Gödel; the consistency of the negation of C.H. relative to Z.F. was demonstrated by Cohen. Thus the present day axioms of set theory are simply not powerful enough to settle the problem.

SECOND SESSION OF THE ASSOCIATION

Session on Mathematical Applications

The Theory of Optimal Growth, by Professor L. W. McKenzie, University of Rochester.

Analogous asymptotic properties of optimal paths of capital accumulation are derived for two models of economic growth, one suggested by the work of John von Neumann and the other by the work of Frank Ramsey. The von Neumann model is specified by a closed convex cone Y in E_{2n}, where $y=(k', -k)\in Y$ and $k', k\geqq 0$. The initial capital stock for a unit period is k and the terminal capital stock is k'. A T period accumulation path is a sequence $\{k^t\}$, where $(k^{t+1}, -k^t)\in Y$ for $t=0, \cdots, T-1$. The objective is to maximize $v(k^T)=\min k_i^T/\bar{k}_i$ for $\bar{k}_i\neq 0$, where $\bar{k}$ is a chosen point in E_n. The Ramsey model is specified by a closed convex set Y in E_{2n+1}, where $y=(u, k', -k)\in Y$ and $u, k', k\geqq 0$. The new variable u represents the level of utility achieved for the consumers during the unit period. The objective is to maximize $\sum_{t=0}^{T} u^t$. Asymptotic results for optimal paths based on a duality lemma due to Roy Radner are proved. Loosely speaking, for an appropriate metric, optimal paths in each model must stay close to a given facet of Y for all but a finite number N of periods, however large T may be.

Geometry and Monotone Functions in Statistics, by Dr. J. B. Kruskal, Bell Telephone Laboratories.

While geometry and monotone functions each enter statistics in areas which are well known to mathematicians, they also enter (sometimes together) in less familiar areas such as "multidimensional scaling," "parametric mapping," "monotone regression," and "factor analysis." The mathematical foundation of these techniques is rather pretty. The use of these techniques often leads to the challenging numerical analysis problem of minimizing a function of very many variables. (This may be illustrated by a brief movie.) In addition, some of these techniques can be very powerful, and provide extremely clear examples of how statistical analysis can provide that elusive thing called "insight."

Mathematical Learning Models for Continuous-Time Parameters and Continuous-Response Space, by Professor P. C. Suppes, Stanford University.

The basic theory of learning models for continuous-time parameters and continuous-response

space was formulated. It was emphasized that the theory depends entirely in its mathematical formulation on the theory of stochastic processes. A survey was given of some of the constructive and testable results that may be obtained. Some of the difficulties of applying the theory to some problems with simple and natural formulation were surveyed.

THIRD SESSION OF THE ASSOCIATION

Hedrick Lecture II, by Professor Fine.

Business Meeting of the Association; presentation of L. R. Ford Awards.

Panel Discussion on the Past, Present, and Future in Mathematics Education, by Dean A. E. Meder, Jr., Rutgers—The State University, Professor E. G. Begle, Stanford University, Professor A. M. Gleason, Harvard University, and Dr. R. S. Fleming, Assistant Commissioner of Education, New Jersey.

Dean Meder opened the discussion by reviewing the past: Although mathematics was completely ignored as an entrance requirement by the American colonial colleges, some expectation of minimal competence in arithmetic developed during the eighteenth century. By the middle of the nineteenth, fairly specific content in algebra and geometry seems to have been generally accepted as standard pre-college study; codification of this content became quite rigid in the first decade of the twentieth century. Three significant quasi-official reform movements developed: one in the early 1920's jointly sponsored by this Association and the National Council of Teachers of Mathematics; a second, embodied in new requirements for the examinations of the College Entrance Examination Board issued in the middle 1930's, which was in many respects a reaction to attacks on mathematics by "educationists," and the most recent and most successful, the work of the Commission on Mathematics in the middle 1950's. In this presentation, these movements were characterized as to their objectives and viewpoints, their degree of success evaluated, and reasons for their successes and failures suggested.

Professor Begle discussed the present: The revolution in mathematics which started over a century ago has now penetrated down into all levels of the school mathematics curriculum. There are now textbooks available for all grade levels which attempt to develop in students not only mathematical skills but also understanding of mathematical ideas and of the structure of mathematical systems. Experience with these new texts has been good on the whole. While some improvements can easily be made, no further major changes are indicated for the near future. The preparation of teachers for these new curricula presents a mixed picture. The recommendations of the Panel on Teacher Training of CUPM for the preparation of secondary school teachers have been generally accepted and are being implemented. The proper training of elementary school teachers, however, seems to be a much more difficult problem.

Professor Gleason then looked at the future: It is agreed that a broad range of mathematical topics will be required by every citizen, and curricula are already being designed to include these new topics. The curricula of the future will differ from the more advanced curricula of today primarily in the order in which the topics are taken up. There will be great emphasis on pre-mathematical experiences, and even the youngest children will be exposed to such concepts as negative numbers, coordinate systems, vectors, symmetry, motion conceived as independent of the object moved, topology, probability and statistics.

Dr. Fleming spoke on the impact of curriculum revision on educational administration. He observed that vigorous activity has been underway in many subject areas of the curriculum since Sputnik. Although the history of American education is rich with illustrations of quests for improvement, schools will continue to seek better opportunities for all young people. The task will never be finished. Activities of recent years have included efforts to up-grade content, provide needed instructional materials, solicit scholars for program development, provide new patterns of in-service education, and to assess the consequence of educational programs. Such activities raise a

series of important questions concerning the role of the teacher, emphasis on *a* book, evaluation, relationships of elementary and secondary school curriculum, use of specialists, and basic purposes to be achieved. Efforts to date seem to be yielding new patterns of teacher education, new sources of materials, and new working relationships with higher education. The current curriculum programs must not become academic issues to be discussed in isolation to teachers and administrators. New relationships are needed with subject area scholars which are cooperative and meaningful. Problems of planning, methodology, evaluation, teacher education and certification are central to the administrator's responsibility.

General Discussion by the Panel and the Audience

FOURTH SESSION OF THE ASSOCIATION

Hedrick Lecture III, by Professor Fine.

Fourier Methods in the Frequency Analysis of Data, by Professor Christopher Bingham, Princeton University and University of Chicago, and Professor J. W. Tukey, Princeton University.

Frequency analysis is an incisive tool in studying the world in spite of such obstacles as complete confounding of every frequency with many others and non-uniqueness of trigonometric expansions of finite data. These techniques have been widely used in geophysics, as such discoveries as the natural periods of vibration of the earth illustrate. Understanding of frequency analysis, its aims and fundamental limitations is founded largely on the formulas of discrete Fourier analysis. The Fast Fourier Transform has made practical the application of better techniques by vastly reducing the effort needed to obtain numerical convolutions and Fourier transforms.

FIFTH SESSION OF THE ASSOCIATION

A Report on the 1966 International Congress of Mathematicians, by Professor Seymour Sherman, Indiana University, Professor J. R. Isbell, Case Institute of Technology, and Professor W. W. Comfort, University of Massachusetts.

Professor Sherman reported that the 1966 International Congress of Mathematicians in Moscow was the largest Congress so far, attended by about 5000 mathematicians from every major country of the world with the exception of Communist China. The opening ceremony and final banquet were held in the beautiful Palace of Congresses in the Kremlin. The Fields medal winners were Michael Atiyah of Oxford University, Paul Cohen of Stanford University, Stephen Smale of the University of California, Berkeley, and Alexandre Grothendieck of Paris. The next Congress is scheduled for September 1970 in Nice, France. The Congress provided a unique opportunity for meeting Russian mathematicians who were present in large numbers and who, in different ways, tried to communicate in their lectures also with those only familiar with western languages. The AMS pretranslated many of the Soviet major addresses, and the Russians rapidly duplicated and distributed these. Some details were provided as to how the speaker, with the help of earlier visitors to the Soviet Union, as well as the judicious use of social lubricants, managed to surmount his linguistic inadequacies and initiate communication with Russian researchers in his field.

Professor Isbell then briefly described the physical plant and functioning of Moscow State University. He commented on the lectures given by Dieudonné, Eršov, S. P. Novikov and Eilenberg.

Professor Comfort finally observed that the courteous and pleasant treatment accorded to Congress members by Russian officials was not always available to Russian citizens under similar circumstances, and told several stories to illustrate this. The Congress was, in general, mercifully free of politics; exceptions to this were the bulletin board at Moscow University protesting "U. S. Atrocities" in Viet Nam, Grothendieck's failure to attend the Congress, Smale's relation with the House Unamerican Activities Committee and his Moscow press conference. Professor Comfort expressed surprise at Russian familiarity with mathematical English-language literature. Statements were given of new important topological results by Lashneff, David Henderson, and Ronald

Knill, as well as detailed statements of advances in cardinal topology by Parovicenko, Juhász, Vopěnka, and Kurepa. He listed pleasant informal social experiences involving Russians, such as swimming in the Moscow River and conversations with non-mathematicians in restaurants, where questions were raised concerning Cassius Clay, Senators Robert and Ted Kennedy, Jack Ruby, etc. He urged American mathematicians to learn Russian, or at least to encourage their graduate students to do so. Periodic perusal of the *Doklady* English-language translations was felt to be insufficient to acquaint an American mathematician thoroughly with Russian work in his field.

SPECIAL SESSIONS OF THE ASSOCIATION

Film showings were held in Room 123 of Scott Hall as follows:

CEM Animated Level I Films (*in color*). Sunday, 7:00–7:14 P.M.—"What is a Set?" 7:15–7:27 P.M.—"One-to-one Correspondence."

CEM Animated Calculus Films (*in color*). Sunday, 7:30–7:50 P.M.—"What is Area?" by Charles E. Rickart; 8:00–8:10 P.M.—"Limit" by Robert C. Fisher; 8:12–8:22 P.M.—"Infinite Acres" by Melvin Henriksen; 8:25–8:34 P.M.—"Fundamental Theorem of the Calculus" by Morris Schreiber; 8:36–8:46 P.M.—"A Function is a Mapping" by Albert Fadell; 8:55–9:05 P.M.—"Area under a Curve" by Charles E. Rickart; 9:07–9:17 P.M.—"The Definite Integral as a Limit" by Robert C. Fisher; 9:19–9:29 P.M.—"Newton's Method" by Herbert S. Wilf; 9:30–9:38 P.M.—"Volume by Shells" by George F. Leger, Jr.

Miscellaneous CEM Films. Monday, 5:00–5:43 P.M.—"Göttingen and New York—Reflections on a Life in Mathematics—Richard Courant" (a CEM Individual Lectures Film in color); 5:45–5:59 P.M.—"What is a Set?" (a repeat showing of the first film listed above); 7:00–7:30 P.M.—"Topology" with Raoul Bott and Marston Morse (A kinescope produced by the *Science and Engineering TV Journal* in b & w); 7:35–8:35 P.M.—"Let us Teach Guessing: A Demonstration with George Polya" (a CEM Individual Lectures Film in color); 8:45–9:15 P.M.—"Mr. Simplex Saves the Aspidistra" with Frank Kocher, Leon Henkin and Julius H. Hlavaty (a CEM Level I film in color). Tuesday, 7:30–8:12 P.M.—"Predicting at Random: A Lecture by David Blackwell" (a CEM Individual Lectures Film in color); 8:20–9:23 P.M.—"Nim, and Other Oriented Graph Games" with Andrew M. Gleason (a CEM Individual Lectures Film in b & w); 9:35–10:40 P.M.—"What is an Integral, New Version" with Edwin Hewitt (a CEM Individual Lectures Film in color).

MEETING OF THE BOARD OF GOVERNORS

The Board of Governors of the Association met on Sunday at 10:00 A.M. in Dining Rooms A, B, and C of the University Commons at Rutgers with thirty-three members present. Among the items of business transacted were the following:

Upon the recommendation of Professor R. A. Rosenbaum, as the incoming Editor of the MONTHLY, the Board elected the following as Associate Editors of the MONTHLY for a five-year term beginning January 1, 1967: Joshua Barlaz, Rutgers—The State University, J. A. Brown, University of Delaware, Leonard Carlitz, Duke University, Howard Eves, University of Maine, Harley Flanders, Purdue University, Raoul Hailpern, SUNY at Buffalo, M. S. Klamkin, Ford Science Laboratory, R. C. Lyndon, University of Michigan, A. P. Mattuck, Massachusetts Institute of Technology, K. O. May, University of Toronto, J. R. Mayor, AAAS, G. N. Raney, University of Connecticut, Gian-Carlo Rota, Rockefeller Institute, E. P. Starke, Rutgers—The State University, J. G. Wendel, University of Michigan. Albert Wilansky, Lehigh University.

The Board voted to invite Professor Gian-Carlo Rota of the Rockefeller Institute to deliver the sixteenth series of Earle Raymond Hedrick Lectures at the 1967 Summer Meeting.

The Board voted to approve the recommendation of the ad hoc Committee on the Graduate Program in Mathematics that the Association become actively interested in the graduate program in mathematics, that it shall concern itself with five problems outlined below, and that CUPM shall be asked to attack these problems, since CUPM already has an office and a staff, and is already known to and respected by mathematicians and administrators. The five problems were listed as follows: (1) Make clear the qualifications appropriate for a college teacher of mathematics at various levels, both as a mathematician and as a teacher. (2) Make qualified teachers recognizable and recognized. (3) Help graduate students go into appropriate programs. (4) Help departments improve (or start) their graduate programs. (5) Hold conferences and summer institutes.

The Board also authorized CUPM to establish a Panel on Mathematics in Two-Year Colleges and specifically requested that it, among its many other duties, concern itself with the problem of building up the interest and activity of this group of mathematics teachers in the MAA.

The Board approved the following schedule of future meetings: Houston, Texas, January 26–28, 1967; University of Toronto, August 28–30, 1967; San Francisco, California, January 25–27, 1968; University of Wisconsin, Madison, August 26–28, 1968; New Orleans, Louisiana, January 1969; University of Oregon, August 25–27, 1969; Miami, Florida, January, 1970.

The Executive Director reported the membership of the Association as 16,940 members, 2 corporate members, and 229 academic members.

The Board voted that, beginning with the fall of 1967, the COMBINED MEMBERSHIP LIST not be sent automatically to all members, as has been the case in the past, but that instead it be sent only to members requesting copies. Forms on which to request free copies of the LIST will be sent to all members in the summer of 1967. Certain classes of members will still receive the LIST automatically including institutional members, family members, members of the Board of Governors, all members of MAA committees, and all Section Officers.

BUSINESS MEETING OF THE ASSOCIATION

A business meeting of the Association was held on Tuesday morning with President Wilder presiding.

The second set of L. R. Ford Awards were presented by President Wilder to authors of expository articles published in the MONTHLY and the MATHEMATICS MAGAZINE in 1965. The awards, in the amount of $100 each, were presented for three articles (for further details on these awards, see the August-September issue of this MONTHLY, page 815).

The Secretary then reported on some of the actions taken by the Board of Governors on Sunday. He announced that the second edition of the *Guidebook to Departments in the Mathematical Sciences in the United States and Canada*, giving information on 1200 institutions, will be ready by October 1 and available from the Buffalo office of the Association for 50 cents per copy. He also announced that the lectures from the 1965 MAA Coöperative Summer Seminar at Bowdoin College on *Harmonic Analysis* by L. H. Loomis are now available from the Buffalo office of the Association for $3.00 per copy postpaid. *Applications of Undergraduate Mathematics in Engineering* by Ben Noble will be published early in 1967. MAA members may purchase single copies at $4.50; nonmembers must purchase copies from Macmillan Company at $9.00 per copy. MAA Studies No. 4, *Studies in Global Geometry and Analysis*, edited by S. S. Chern, will be published early in 1967.

The Secretary, speaking in behalf of all those who attended the meeting, extended the thanks of the Association to Rutgers for having made their facilities available for this

meeting and for all the assistance which has been received from the members of the department of mathematics and the other staff at Rutgers.

MEETING OF SECTION OFFICERS

The meeting of representatives of the Sections was held on Monday evening in Dining Rooms A, B, and C of the University Commons. Professor L. E. Mehlenbacher, Chairman of the Committee on Sections, presided. Fifty-seven persons were present, representing twenty-six of the twenty-seven Sections of the Association.

President R. L. Wilder discussed the regional educational laboratories which have just been established in various parts of the country, and urged that scientists and mathematicians, in particular, contact these laboratories to offer their assistance so that their viewpoints are properly represented.

Professor H. M. Gehman spoke about some of the problems of the Buffalo office in trying to keep accurate records of members' addresses. In particular, he requested Section Secretaries to notify the Buffalo office of any errors in address lists, of deaths of MAA members, and changes of names of institutions.

Professor H. L. Alder announced that arrangements between Modern Learning Aids (MLA) and the MAA provide for free showing of the films produced by the Individual Lectures Project of CEM at sectional and national meetings of the MAA. Section Officers desiring to show such films at sectional meetings should write to MLA at one of its branches or its headquarters, 1212 Avenue of the Americas, New York, New York, 10036, and specify that the showing will be at a sectional meeting. The Level I films shown at this meeting are expected to be available from MLA as of January 1, 1967. Arrangements for general distribution of the calculus films have not as yet been completed, but they are available for showing at sectional meetings from Professor H. M. MacNeille, Department of Mathematics, Case Institute of Technology, University Circle, Cleveland, Ohio, 44106.

Professor Alder then appealed to the Section Officers to see to it that they have a strong local organization and that they be conscious of their responsibility to carry out at the Section level the many projects initiated by the MAA at the national level. He also urged a strengthening of the system of MAA representatives. It is the responsibility of the Section Chairmen to appoint at each college and university within their Section an MAA representative. He suggested that, in larger departments, MAA representatives write circular letters to their colleagues to inform them of current MAA activities of interest to them, such as CUPM recommendations, newly available mathematical films, and coming sectional and national meetings of the MAA. He also urged MAA representatives to take the initiative in inviting MAA Visiting Lecturers to their campuses and to arrange for showings of films produced by the MAA.

Professor K. O. May, Chairman of the Committee on the Preparation of a Fifty-Year History of the Association, announced that his Committee had decided to include a history of the Sections of the Association and, therefore, requested that each Section send him information on its history, especially its unique features, and also supply him with names of old members for possible interviews.

Professor R. D. Larsson, Chairman of the Upper New York Section, reported on the Undergraduate Paper Contest of that Section, which was established during the 1965–66 academic year. Department chairmen were invited to nominate papers prepared by undergraduate students in their institutions. Five papers were nominated, and one was chosen for presentation at the spring meeting of the Section. Prizes of \$25, \$15 and \$10 were awarded to the top three papers. This money was made available to the Section by the Association. The success of the contest has resulted in its extension for at least another year.

Professor C. T. Salkind, Chairman of the Committee on High School Contests, gave a

report on correlation studies between the MAA High School Contest and others. He emphasized that to date no quantitative or statistical study has been undertaken, chiefly because it does not seem that any useful purpose would be served by such a study. Since the objectives of the High School Contest and such other examinations as the SAT or CEEB are significantly different, one would be reasonably safe in predicting a low correlation.

Professor S. A. Jennings led a discussion on "How can we involve members of mathematics departments in community colleges and junior colleges in the activities of the MAA?" He reported on a panel discussion on mathematics in two-year institutions which he had organized in the Pacific Northwest Section two years ago for a Section meeting. As a result, a committee of two-year college mathematicians was encouraged to meet jointly with the Pacific Northwest Section at their regular June meeting and have a program of their own choosing.

Professor G. S. Young led a discussion on "How can we improve the nature of Section programs?" He expressed his feeling that there is a very large number of papers presented at Section meetings which can be described as very minor research papers and that presentation of such papers will have a most damaging effect on interest in Section meetings. Various suggestions were offered for avoiding this: (1) by having parallel sessions at Section meetings with expository papers and contributed research papers being scheduled simultaneously, (2) by meeting jointly with the AMS, as is done in the Pacific Northwest where research papers then are presented at the AMS sessions and all papers at the MAA session are presented by invitation only, and (3) by a careful selection of topics and speakers supervised by a program committee appointed by the Section for this purpose.

MEETINGS OF OTHER ORGANIZATIONS

The American Mathematical Society held its session from Tuesday afternoon through Friday noon. Invited addresses were given by Professor Nachman Aronszajn of the University of Kansas on "Miscellanea on Hilbert Spaces, Linear Operators, and Numerical Ranges" on Tuesday at 2:00 P.M.; by Professor H. J. Keisler of the University of Wisconsin on "Recent Developments in Model Theory" on Wednesday at 4:00 P.M.; by Professor V. S. Varadarajan of the University of California, Los Angeles, on "Representations of Semisimple Lie Groups" on Thursday at 9:00 A.M.; and by Professor Karl Menger of the Illinois Institute of Technology on "Ideas on Complex Functions" on Friday at 9:00 A.M.

The Institute of Mathematical Statistics held sessions from Tuesday afternoon through Friday. A special invited paper was presented by Professor David Friedman of the University of California, Berkeley, on "Some Invariance Principles for Functionals of a Markov Chain" on Wednesday at 10:50 A.M., and an invited address by Professor R. E. Kalaba of the RAND Corporation on "Mathematical Challenges of Biological Processes" on Thursday at 11:30 A.M.

The Society for Industrial and Applied Mathematics presented the von Neumann Lecture at 8:00 P.M. on Wednesday. It was delivered by Professor E. P. Wigner of Princeton University on "Statistical Theory of Spectra."

The Pi Mu Epsilon Fraternity held sessions for contributed papers on Tuesday at 3:00 P.M. and on Wednesday at 9:00 A.M. A banquet was held Tuesday evening at 6:00 P.M. in Dining Rooms A, B, and C of the University Commons. At the banquet, Professor J. S. Frame, retiring national president of Pi Mu Epsilon, spoke on "Semi-inverses of a Matrix." A Dutch-treat breakfast meeting for Pi Mu Epsilon members was held on Wednesday at 8:00 A.M.

The Governing Council of Mu Alpha Theta, the National High School and Junior College Mathematics Club, sponsored by the MAA, met on Wednesday at 9:00 A.M. in Room 4 on Level B of Frelinghuysen Hall.

ARRANGEMENTS, ENTERTAINMENT AND RECREATION

The Committee on Arrangements for the meeting consisted of L. F. McAuley, Chairman; H. L. Alder, Joshua Barlaz, John Bender, Saul Blumenthal, F. E. Clark, R. M. Cohn, J. H. Griesmer, Katherine E. Hazard, Mrs. Barbara L. Osofsky, Everett Pitcher, M. S. Robertson, G. L. Walker, and K. G. Wolfson.

Registration headquarters were located in The Ledge. Dormitory accommodations and cafeteria facilities were provided by Rutgers. The Mathematical Sciences Employment Register was maintained on Level A of Frelinghuysen Hall, and book exhibits, exhibits of educational media and a group of drawings and paintings of eleven great mathematicians of the past, made by Professor Hwa S. Hahn of Pennsylvania State University, were displayed in Records Hall.

There was a chicken fry on Tuesday at 6:00 P.M. on the mall in front of William the Silent, followed by a square dance at 8:00 P.M. in the Gymnasium parking lot. There was a Presidential Tea on Wednesday at 4:00 P.M. on the President's lawn. SIAM conducted a beer party on Monday at 8:30 P.M. in the Officers' Club at Camp Kilmer, and IMS a mixer on Wednesday at 9:00 P.M. in the University Commons. Non-mathematical movies were shown Monday through Friday at 8:30 P M. in Milledoler Hall, Room 100, at no charge.

HENRY L. ALDER, *Secretary*

ACADEMIC MEMBERS ELECTED INTO THE ASSOCIATION

In accordance with the amendments adopted at the business meeting of the Association at Stillwater on August 30, 1961, the Board of Governors at its meeting at Rutgers—The State University, New Brunswick, New Jersey, on August 28, 1966, elected to membership in the Association the tenth set of applicants for academic members (for election of the other nine sets, see the April and November issues for 1962–66). Approval for election to membership was given to the following 7 applicants for academic membership:

Kansas State College
Memorial University of Newfoundland
New Haven College
Sacred Heart University
St. Mary's College of Maryland
State University College, Cortland, New York
U. S. Naval Academy, Annapolis

HENRY L. ALDER, *Secretary*

RETIRED MATHEMATICIANS

The next issue of the List of Retired Mathematicians Available for Employment will be published in January 1967 by the Mathematical Sciences Employment Register. Besides being distributed to academic and industrial employers, who request it from the Register office, copies will be available when the Register is open at the annual meeting in Houston, Texas, January 25–27, 1967. Retired mathematicians who are interested in being included in the list may either request a form from the Register office or send the following information: name, date of birth, highest degree earned and where it was obtained, most recent employment, present address, date available, references, preference for academic or industrial employment, and geographic location preferred. The deadline for receipt of either the completed form or the above information is January 1, 1967.

SUMMER EMPLOYMENT OPPORTUNITIES

The List of Opportunities for Summer Employment for Mathematicians and College Mathematics Students is being compiled and will be issued in January 1967. The list will be available free of charge at the annual meeting in Houston, Texas, in January and on request from the Register office. Institutions and industrial concerns that have summer openings and would welcome applications from mathematicians and students of mathematics may request forms for listing from the Register office. The deadline for receipt of completed forms is January 1, 1967.

THE EMPLOYMENT REGISTER

The Mathematical Sciences Employment Register, established by the American Mathematical Society, the Mathematical Association of America, and the Society for Industrial and Applied Mathematics, will be open at the annual meeting at the Rice Hotel in Houston, Texas. The dates are January 25–27, 1967. The Register will be open from 9 A.M. to 5 P.M. on Wednesday, Thursday, and Friday in the Terrace Room. Applicants and employers who wish to be listed should request forms from the Register office. Deadline for receipt of completed forms is December 15, 1966. Employers who enter listings after December 15 are charged $5.00 for each late listing. If these listings arrive too late to be included in the printed issues, they will be taken to the meeting and will be on display for those interested in seeing them.

There is no charge for listing except when the late registration fee for employers is applicable. Provision will be made for anonymity of applicants upon request and upon payment of $5.00 to defray the cost involved in handling anonymous listings.

A subscription, which includes three issues of the List of Applicants and the List of Positions is $15.00 a year. These lists are published in January, May, and August. The January issues will be mailed the first week of January. Individual sets of the lists are $7.50, and copies of the List of Positions only may be purchased for $3.00. Copies of the lists will be available at the January meeting or may be ordered in advance from the Register office. Checks should be made payable to the American Mathematical Society.

SECOND EDITION OF GUIDEBOOK

MAA has just published a second edition (1966) of the GUIDEBOOK TO DEPARTMENTS IN THE MATHEMATICAL SCIENCES. This new edition consists of about eighty pages, and contains almost twelve hundred entries. Copies may be purchased at fifty cents each from the Buffalo office of MAA.

The GUIDEBOOK, edited by Professor Raoul Hailpern, is a project of the MAA Committee on Advisement and Personnel. It is intended to provide in summary form information about the location, size, staff, library facilities, course offerings, and special features of departments in the mathematical sciences in four year colleges and universities in the United States and Canada. Its purpose is to assist prospective students and their counselors and parents in obtaining comparable information about many institutions so that the selection of a proposed place of study may be narrowed down to a few from which more detailed information may then be sought.

The GUIDEBOOK gives data for institutions that offer an undergraduate major in a mathematical science, such as Mathematics, Statistics, Computer Science, Applied Mathematics, leading to a baccalaureate degree. Additional information is given for institutions that offer the Ph.D. degree in one or more mathematical sciences. This information concerns the size, financial support and productivity of the doctoral programs of these institutions.

MAY MEETING OF THE NEW JERSEY SECTION

The first joint meeting of the New Jersey Section of the MAA and the Association of Mathematics Teachers of New Jersey was held at Ocean Township High School, Oakhurst, on May 7, 1966. During the morning program, Professor Joshua Barlaz, Chairman of the Section, presided at the general session; Professor F. A. Varrichio, Secretary-Treasurer, presided at the elementary and junior high school session; and Professor F. E. Clark, Vice-President presided at the senior high school and college session. Mrs. Margret H. Devitt, President of the AMTNJ, presided at the luncheon session. One hundred and eighty-five persons attended the meeting, including sixty-six members of MAA.

There was no business meeting.

At the general session, the following paper was presented:

Geometric concepts from Euclid to Einstein, by A. W. Tucker, Princeton University (by invitation).

At the senior high school and college session, the following paper was presented:

Statistics—an enigma, by E. R. Ott, Rutgers—The State University (by invitation).

At the elementary and junior high school session, the following paper was presented:

The emerging program in mathematics for grades 5–8, by W. G. Koellner, Coordinator of Mathematics, Cranford Public Schools (by invitation).

The program in mathematics, grades 5–8, is relatively independent of the grade level structure of the school system. Grade level structures, such as K-6-3-3, K-6-4-2, K-5-4-3, and K-4-4-4, are more dependent on socio-economic conditions and administrative viability than on specific subject content or offerings.

The report of the Cambridge Group projects directions and long-range goals for school mathematics. Investigations into current programs indicate that certain of these goals are being realized. Research indicates that these goals are attainable and necessary.

Increased emphasis must be placed on training teachers in content as well as method. Inservice training is necessary to develop depth and breadth of knowledge.

At the luncheon session the following paper was presented:

Mathematical education for scientific, technological and industrial needs of society, by H. F. Fehr, Teachers College, Columbia University (by invitation).

Present mathematical reform in school is limited to updating the traditional program. Subject separation is maintained. For efficiency and real learning, mathematics must be conceived as a totality. The high school program must be reconstructed, using sets, mappings, relations and functions as unifying concepts. On these ideas are built the properties of the structures of number systems, groups, rings, fields—all headed toward vector spaces. Arithmetic, algebra, geometry, and analysis are completely unified. To culminate the secondary school program, probability including statistical inference, the calculus, and numerical analysis related to digital computers, should be studied.

F. A. VARRICHIO, *Secretary-Treasurer*

MAY MEETING OF THE UPPER NEW YORK STATE SECTION

The annual meeting of the Upper New York State Section of the MAA was held at St. Bonaventure University, Olean, New York, on Saturday, May 14, 1966. M. W. Pownall, Chairman of the section, presided at the morning session, and R. D. Larsson,

Vice-Chairman, presided at the afternoon session. There were 135 persons in attendance, including 104 members of the association.

At the business meeting the following officers were elected: Chairman, R. D. Larsson, Mohawk Valley Community College; Vice-Chairman, Dick Wick Hall, SUNY at Binghamton; Secretary-Treasurer, N. G. Gunderson, University of Rochester. R. D. Larsson, chairman of the Undergraduate Paper Contest Committee, reported the successful launching of this new venture of the section. J. M. Perry, chairman of the Committee to Study the CUPM Report, *General Curriculum in Mathematics for Colleges*, called for reaction to the report to be sent to his committee. The Section moved its congratulations to Professor H. M. Gehman on his receiving the Award for Distinguished Service in Mathematics. B. H. Gere presented the findings of the Committee on Time of Meeting, and moved that in addition to the annual meeting in the spring, fall meetings be held for two years on an experimental basis. D. S. Martin reported on a survey of mathematics club activities in the section. Nura Turner reported on the 1966 MAA High School Mathematics Contest. N. G. Gunderson reported that section membership now exceeds 1000, and increased expenses make it necessary to levy a nominal registration fee at future meetings.

The program was as follows:

1. *Trends in undergraduate programs for mathematics majors*, a panel discussion by M. W. Pownall, Colgate University; R. Exner, Syracuse University; J. M. Perry, Wells College; R. D. Slaugh, SUNY at Cortland; A. D. Zieber, SUNY at Binghamton.

Programs at the colleges represented were described. Trends include the reduction in the amount of analytic geometry in the calculus course, linear algebra before multivariate calculus, no separate course in differential equations, real analysis instead of advanced calculus, the requirement of 6 hours of mathematics for elementary school teachers, honors programs and independent study.

2. *The general curriculum in mathematics for colleges*, by A. W. Tucker, Princeton University, representing CUPM.

3. *A generalization of the consecutive non-square free integer problem*, by C. J. Parry, SUNY at Oswego, Undergraduate Paper Contest winner.

It was shown that for all positive integers n and k there exist n consecutive integers each with a kth power factor.

4. *Present and future trends of mathematics programs in two-year colleges*, a panel discussion by R. D. Larsson, Mohawk Valley Community College; Mrs. R. Humphrey, Bennett College; C. A. Lathan, Monroe Community College; J. J. Deveny, SUNY at Morrisville.

The typical mathematics programs of the three types of two year colleges represented were presented, with some development of the historical patterns. Emphasis was given on the variety of curricula to be served; the spread of backgrounds of the students; the conflict between career and transfer objectives; the lack of adequate texts and the need and opportunity for new curricular development in this area.

Discussion was held on the special problems of career programs, transfer programs in Arts and Sciences, remedial courses and the objectives of mathematics courses for those not majoring in mathematics or science.

5. *Secondary school mathematics—present and future*, by Frank Hawthorne, N. Y. State Education Department.

6. *Finite calculus*, by H. A. Still, Queen's University.

It would be convenient in the study of a series if we could find the sum of the first n terms of a series, that is to obtain a simple expression for the partial sums, S_n, of a series. The difference

operator Δ was defined for functions whose domain is the natural numbers. This was followed by a definition of the operators Δ^{-1} and $\sum_a^b$. The words differentiable and summable were defined and the following Fundamental Theorem was presented: If f is summable on $\{x | a \leqq x \leqq b\}$ and F is such that $\Delta F(x) = f(x)$, then $\sum_a^b f(x) = F(b+1) - F(a)$.

7. *Projection matrices and the generalized inverse*, by J. C. Boot, SUNY at Buffalo.

It is shown that the generalized inverse can be considered as the matrix X which is such that, out of all matrices X that minimize $\mathrm{tr}(AX-I)'(AX-I)$ (that is in words: all matrices X such that AX is as close in a Euclidean sense to I as possible) it is that matrix for which tr $X'X$ is minimized. This leads to the elegant formula for the generalized inverse $X = A'A_r(A_r'AA'A_r)^{-1}A_r'$, A of order $n \times m$ and A_r any r independent columns of A.

N. G. GUNDERSON, *Secretary*

JUNE MEETING OF THE PACIFIC NORTHWEST SECTION

The annual meeting of the Pacific Northwest Section of the MAA was held at the University of Victoria, Victoria, British Columbia on June 17 and 18, 1966. There were 179 persons in attendance including 85 members of the Association. Professor S. A. Jennings, Chairman of the Section, presided over the Friday sessions which were held jointly with the Northwest Section of SIAM. Mr. Joel Carlson presided over the Saturday sessions which were devoted to junior and community colleges. Those in attendance were guests of the Departments of Mathematics of the University of Victoria and the Canadian Services College at a social gathering on Friday evening and were guests of the Centennial Committee of the Province of British Columbia at a luncheon on Saturday. A reception and tea was held at the Canadian Services College at the conclusion of the meeting.

At the business meeting the following officers were elected: Chairman, W. R. Ballard, Montana State University; Vice-Chairman, C. T. Long, Washington State University; Secretary-Treasurer, E. A. Maier, University of Oregon.

The program of the joint sessions with SIAM was as follows:

1. *On uniqueness in the mathematical theory of plasticity*, by E. M. Shoemaker, Simon Fraser University (by invitation).

2. *Panel discussion: The CUPM report on a general curriculum in mathematics;* H. E. Chrestenson, Reed College; Ivan Niven, University of Oregon; T. M. Southward, California State College at Hayward; Roy Dubisch, University of Washington (Chairman).

3. *Boolean Algebras—some old and new problems*, by R. S. Pierce, University of Washington (by invitation).

4. *An application of Kalman filtering to passive ranging*, by J. W. Burrows, The Boeing Company (by invitation).

The following program was presented at the junior and community college sessions:

1. *Comments concerning the two-year colleges of the Pacific Northwest*, by S. A. Jennings, University of Victoria.

2. *Is there a place for the two-year college in the NCTM?* by Donovan Johnson, President, NCTM, University of Minnesota.

3. *Is there a place for the two-year college in the MAA?* by R. C. James, Member, CUPM, Harvey Mudd College.

4. *Panel discussion: Mathematics at the two-year college—points of view;* Glenn Adams, Everett Junior College; Dana Lefstad, Skagit Valley College; Arvid Lonseth, Oregon State University; Brice Whittles, British Columbia Institute of Technology and Selkirk College.

E. A. MAIER, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

Forty-eighth Summer Meeting, University of Toronto, Toronto, Ontario, Canada, August 28–30, 1967.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Associate Secretary.

ALLEGHENY MOUNTAIN, West Virginia University, Morgantown, West Virginia, May 6, 1967.
ILLINOIS, University of Illinois, Urbana, May 14–15, 1967.
INDIANA
IOWA, Drake University, Des Moines, April 21, 1967.
KANSAS, Fort Hays State College, Hays, April 22, 1967.
KENTUCKY, Murray State University, Murray, Spring, 1967.
LOUISIANA-MISSISSIPPI, Jung Hotel, New Orleans, Louisiana, March 4–5, 1967.
MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Georgetown University, Washington, D.C., December 10, 1966.
METROPOLITAN NEW YORK
MICHIGAN
MINNESOTA
MISSOURI, Northeast Missouri State Teachers College, Kirksville, April 29, 1967.
NEBRASKA, University of South Dakota, Vermillion, May 6, 1967.
NEW JERSEY
NORTHEASTERN
NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.
OHIO
OKLAHOMA-ARKANSAS, Northeastern State College, Tahlequah, Oklahoma, March–April, 1967.
PACIFIC NORTHWEST, University of Montana, Missoula, June 16–17, 1967.
PHILADELPHIA
ROCKY MOUNTAIN
SOUTHEASTERN, Florida Presbyterian College, St. Petersburg, Florida, March 31–April 1, 1967.
SOUTHERN CALIFORNIA, San Diego State College, San Diego, March 11, 1967.
SOUTHWESTERN, University of Arizona, Tucson, March 31–April 1, 1967.
TEXAS, Austin College, Sherman, April 14–15, 1967.
UPPER NEW YORK STATE, State University College, Plattsburgh, May 20, 1967.
WISCONSIN, St. Norbert College, DePere, May 6, 1967.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, D. C., December 26–31, 1966.
AMERICAN MATHEMATICAL SOCIETY, Houston, Texas, January 24–27, 1967.
AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Michigan State University, June 19–23, 1967.
ASSOCIATION FOR COMPUTING MACHINERY, Sheraton-Park, Washington, D. C., August 29–31, 1967.
ASSOCIATION FOR SYMBOLIC LOGIC, Houston, Texas, January 23–24, 1967.
CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Chicago, November 23–25, 1967.
INSTITUTE OF MATHEMATICAL STATISTICS
NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Houston, Texas, January 28, 1967.
OPERATIONS RESEARCH SOCIETY OF AMERICA, New York Hilton Hotel, May 31–June 2, 1967.
SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS

THE AMERICAN MATHEMATICAL MONTHLY, Vol. 73, No. 9, November, 1966, pp. 937-1048

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN

INDEX TO VOLUME 73, 1966

THE AMERICAN MATHEMATICAL MONTHLY

GENERAL MATHEMATICAL PAPERS

ALGEBRA AND THEORY OF NUMBERS

ANALYSIS

GEOMETRY AND TOPOLOGY

PROBABILITY AND STATISTICS

OTHER APPLICATIONS

PROFESSIONAL MATTERS

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS and University of Maryland

COLLABORATING EDITORS: JOHN D. BAUM, Oberlin College and JOHN A. BROWN, University of Delaware

BRIEF COMMENTS

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE

COLLABORATING EDITORS: J. BARLAZ, Rutgers—The State University; L. CARLITZ, Duke University; H. S. M. COXETER, University of Toronto; H. EVES, University of Maine; M. S. KLAMKIN, Ford Scientific Laboratory; A. E. LIVINGSTON, University of Alberta; and A. WILANSKY, Lehigh University.

PROBLEMS PROPOSED

PROBLEMS SOLVED

SOLUTIONS

Numbers in **boldface** type refer to problems, those in lightface, to pages.

MISCELLANEOUS

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University

COLLABORATING EDITORS: KENNETH O. MAY, University of Toronto and E. P. VANCE, Oberlin College

BRIEF MENTION

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

PERSONAL ITEMS

GENERAL INFORMATION

NECROLOGY

REPORTS AND ANNOUNCEMENTS OF THE ASSOCIATION AND ITS SECTIONS

MEETINGS AND ANNOUNCEMENTS OF THE ASSOCIATION

MEETINGS OF ITS SECTIONS

PERSONAL INFORMATION

The following persons presented papers at meetings of the Association and its Section:

THE AMERICAN MATHEMATICAL MONTHLY

THE OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA, INC.

VOLUME 73 NUMBER 10

CONTENTS

DECEMBER 1966

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

FREDERICK A. FICKEN, *Editor*

EDITORIAL CORRESPONDENCE should be addressed to the Editor, F. A. Ficken, Department of Mathematics, New York University, New York, N. Y. 10453. *Manuscripts of papers should be submitted to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.*

ADVERTISING CORRESPONDENCE should be addressed to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

NOTICE OF CHANGE OF ADDRESS by members of the Association as well as correspondence regarding subscriptions to the MONTHLY should be sent to the Executive Director, H. M. GEHMAN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N. Y. 14214.

THIS IS THE OFFICIAL JOURNAL OF THE

MATHEMATICAL ASSOCIATION OF AMERICA, INC.

(Devoted to the Interests of Collegiate Mathematics)

General Office: SUNY at Buffalo, Buffalo, N. Y. 14214. *Executive Director:* H. M. GEHMAN

Annual dues for members of the Association (including a subscription to the American Mathematical Monthly) are $6.00. For nonmembers the subscription price is $10.00.

PUBLISHED BY THE ASSOCIATION at Menasha, Wisconsin, and Buffalo, N. Y. during the months of January, February, March, April, May, June–July, August–September, October, November, December.

Second-class postage paid at Menasha, Wisconsin. Acceptance for mailing at special rate of postage provided for in the Act of February 28, 1925, embodied in Paragraph 4, Section 538, P. L. and R., authorized April 1, 1926.

THE DEFINITION OF A QUADRATIC FORM

ANDREW M. GLEASON, Harvard University

Let V be a vector space over a field F. We exclude, once and for all, the possibility that F have characteristic 2, but otherwise F may be arbitrary. The vector space V may have finite or infinite dimension.

A quadratic form on V is classically defined as a function Q from V to F which can be expressed in the form

$$Q(v) = \sum_{\alpha,\beta} \lambda_{\alpha,\beta} f_\alpha(v) f_\beta(v),$$

where $\{f_\alpha\}$ is the family of linear functionals dual to some basis $\{x_\alpha | \alpha \in A\}$ of V and $\{\lambda_{\alpha,\beta}\}$ is any family of constants indexed on $A \times A$. For each v this sum will be effectively finite. A more intrinsic definition describes a quadratic form as the composition of some bilinear form from $V \times V$ to F with the diagonal injection of V into $V \times V$; that is,

$$Q(v) = B(v, v),$$

where B is bilinear.

Each bilinear form thus defines a quadratic form, but the correspondence is not one-to-one since B and the symmetric bilinear form S defined by

$$S(v, w) = \tfrac{1}{2}(B(v, w) + B(w, v))$$

define the same quadratic form. On the other hand, there is a one-to-one correspondence between symmetric bilinear forms and quadratic forms since we can recover S from Q by the formula

$$4S(v, w) = Q(v + w) - Q(v - w). \tag{1}$$

The study of quadratic forms is thereby reduced to the study of symmetric bilinear forms.

There is something inappropriate about defining a quadratic form which is a function of one variable, in terms of a bilinear form which involves two variables. This raises the question of what requirements can be imposed on a function from V to F to define the set of all quadratic forms.

The best known identity satisfied by quadratic forms is the parallelogram law

$$Q(v + w) + Q(v - w) = 2Q(v) + 2Q(w). \tag{2}$$

For the convenience of the reader, we reprove the theorem of von Neumann and Jordan [**3**] concerning the parallelogram law.

0.1 THEOREM. *Let Q be a function from V to F which satisfies* (2). *The function S defined by* (1) *is symmetric and biadditive and* $Q(v) = S(v, v)$.

Proof. Putting $v = w = 0$ in (2), we see that $Q(0) = 0$. Then putting $v = 0$, we find that $Q(-w) = Q(w)$. From this it follows that S is symmetric. Finally,

putting $v=w$, we get $Q(2v)=4Q(v)$ and derive $S(v, v)=Q(v)$.

Since S is symmetric, to prove biadditivity it will be sufficient to prove it is additive in its first argument. Let x, y, and z be any three vectors in V.

$$\begin{aligned} 8S(x,z) + 8S(y,z) &= 2Q(x+z) + 2Q(y+z) - 2Q(x-z) - 2Q(y-z) \\ &= Q(x+y+2z) + Q(x-y) - Q(x+y-2z) - Q(x-y) \\ &= 4S(x+y, 2z), \end{aligned}$$

whence

$$S(x, z) + S(y, z) = \tfrac{1}{2}S(x + y, 2z). \tag{3}$$

Consider the special case $y=0$. Using $S(0, z)=S(z, 0)=0$ (direct from (1)), we obtain the identity $S(x, z)=\frac{1}{2}S(x, 2z)$. Hence (3) becomes

$$S(x, z) + S(y, z) = S(x + y, z).$$

0.2 COROLLARY. *Let V be a vector space over F of dimension at least two. If Q is a function from V to F such that the restriction of Q to each two dimensional subspace of V is a quadratic form, then Q is a quadratic form.*

Proof. Let v and w be any two vectors in V. Since the restriction of Q to the linear subspace spanned by v and w is a quadratic form the parallelogram identity (2) must hold. Hence $Q(v)=S(v, v)$ where S is a symmetric biadditive function from $V\times V$ to F. To check that S is homogeneous, i.e., that $S(\lambda v, w)=\lambda S(v, w)$, for a fixed v and w requires only reference to the subspace spanned by v and w. Since Q is a quadratic form on this subspace, S is homogeneous.

The corollary does not help us to find a suitable definition of a quadratic form. As is well known, additivity implies homogeneity with respect to scalars in the prime subfield. However, the parallelogram identity cannot possibly imply the homogeneity of S when F is not a prime field. In that case there exists an additive map θ of F into itself which is not homogeneous over F; then if Q is a nontrivial quadratic form, $\theta \circ Q$ satisfies (2) but the corresponding biadditive form $\theta \circ S$ is not homogeneous. Obviously some identity involving the scalar multiplication is necessary to define a quadratic form.

While teaching an elementary course in which the only field involved was the real numbers, I conjectured that the parallelogram law together with the identity

$$Q(\lambda x) = \lambda^2 Q(x) \tag{4}$$

would suffice. This seemed plausible since almost any regularity assumption in conjunction with additivity suffices to guarantee linearity for functions from the reals to the reals. However, the conjecture is false. For fields of characteristic 0, the identities (2) and (4) guarantee the homogeneity of S with respect to algebraic scalars, but not with respect to transcendental scalars!

In this paper we shall prove this and discuss some other identities satisfied by quadratic forms.

1. Quasi-quadratic forms. We shall denote by $\mathfrak{Q}$ the set of all quadratic forms defined on V, and by $\mathfrak{Q}_0$ the set of all functions from V to F which satisfy the identities (2) and (4). Evidently $\mathfrak{Q}_0$ is a linear subspace of the space of all functions from V to F and $\mathfrak{Q}$ is a linear subspace of $\mathfrak{Q}_0$. We shall refer to a member of $\mathfrak{Q}_0$ as a *quasi-quadratic* form.

We associate with each quasi-quadratic form Q a symmetric biadditive form S using the definition (1) and a function E from $F\times V\times V$ to F defined by

$$E(\lambda, v, w) = S(\lambda v, w) - \lambda S(v, w). \tag{5}$$

It is obvious that E is additive in each of its three arguments. Furthermore, it is clear that E is identically zero if and only if S is actually linear in its first argument. Since S is symmetric, this amounts to the assertion that E is zero if and only if $Q\in\mathfrak{Q}$, i.e., Q is a genuine quadratic form.

Let $\mathcal{E}$ be the set of all functions E from $F\times V\times V$ to F which are

(a) linear and skew-symmetric in their second and third arguments, and

(b) additive derivations in their first argument, i.e.,

$$E(\lambda + \mu, v, w) = E(\lambda, v, w) + E(\mu, v, w) \quad \text{and}$$

$$E(\lambda\mu, v, w) = \lambda E(\mu, v, w) + \mu E(\lambda, v, w).$$

The set of such functions is obviously a linear space over F.

1.1 THEOREM. *The map Θ: $Q\to E$ defined by* (1) *and* (5) *is a linear surjection from $\mathfrak{Q}_0$ to $\mathcal{E}$ with kernel $\mathfrak{Q}$.*

Proof. Clearly Θ is a linear map from $\mathfrak{Q}_0$ to the set of all functions from $F\times V\times V$ to F. We have already seen that its kernel is $\mathfrak{Q}$. The essential step is to show that E, as defined in (1) and (5), is actually in $\mathcal{E}$.

It follows directly from (1) and (2) that $Q(x)+2S(x, y)+Q(y)=Q(x+y)$. Replace y by λx, expand using (4), and cancel terms to get

$$S(\lambda x, x) = \lambda S(x, x).$$

Polarize the latter: Replace x by $x+y$, expand by biadditivity and cancel terms.

$$S(\lambda x, y) + S(\lambda y, x) = \lambda(S(x, y) + S(y, x)). \tag{6}$$

Transposing terms, we obtain $E(\lambda, x, y)=-E(\lambda, y, x)$. Thus E is skew-symmetric in its second and third arguments.

Replace x by λx in (6). Since $S(\lambda x, \lambda y)=\lambda^2S(x, y)$ by (1) and (4), we get

$$S(\lambda^2x, y) + \lambda^2S(x, y) = 2\lambda S(\lambda x, y).$$

Polarize again, replacing λ by $\lambda+\mu$, etc. Recall that S is homogeneous with respect to the scalar 2.

$$S(\lambda\mu x, y) + \lambda\mu S(x, y) = \lambda S(\mu x, y) + \mu S(\lambda x, y). \tag{7}$$

Transposing terms, we find $E(\lambda, \mu x, y)=\mu E(\lambda, x, y)$, that is, E is homogeneous

in its second argument. By skew symmetry it is also homogeneous in its third argument. This shows that E satisfies condition (a) in the definition of $\mathcal{E}$. Subtracting $2\lambda\mu S(x, y)$ from both sides of (7), we find that E also satisfies condition (b). Hence $E \in \mathcal{E}$.

Finally we must show that Θ is surjective. Suppose $E_0 \in \mathcal{E}$. We will construct a quasi-quadratic form Q with $\Theta Q = E_0$.

Choose a basis $\{x_\alpha | \alpha \in A\}$ of V and let $\{f_\alpha\}$ be the dual family of linear functionals on V. Then we have $v = \sum_\alpha f_\alpha(v) x_\alpha$ for any $v \in V$, the sum being effectively finite.

Define a function B from $V \times V$ to F by

$$B(v, w) = \sum_\alpha E_0(f_\alpha(v), x_\alpha, w)$$

and put $Q(v) = 2B(v, v)$. Since B is additive in its first argument and linear in its second, it follows immediately that Q satisfies the parallelogram identity (2). Moreover

$$\begin{aligned} Q(\lambda v) &= 2B(\lambda v, \lambda v) = 2\lambda B(\lambda v, v) \\ &= 2\lambda \sum_\alpha E_0(\lambda f_\alpha(v), x_\alpha, v) \\ &= 2\lambda \sum_\alpha \lambda E_0(f_\alpha(v), x_\alpha, v) + 2\lambda \sum_\alpha f_\alpha(v) E_0(\lambda, x_\alpha, v) \end{aligned}$$

since E_0 is a derivation in its first argument. Since E_0 is linear in its second argument and skew symmetric in its second and third, the second sum is

$$\sum_\alpha E_0(\lambda, f_\alpha(v) x_\alpha, v) = E_0(\lambda, v, v) = 0.$$

Thus $Q(\lambda v) = \lambda^2 Q(v)$ and Q is a quasi-quadratic form.

The symmetric biadditive form S associated with Q is given by

$$S(v, w) = B(v, w) + B(w, v)$$

and $\Theta Q = E$, where $E(\lambda, v, w) = B(\lambda v, w) - \lambda B(v, w) + B(w, \lambda v) - \lambda B(w, v)$. Since B is linear in its second argument, the last two terms cancel. Hence,

$$\begin{aligned} E(\lambda, v, w) &= \sum_\alpha [E_0(\lambda f_\alpha(v), x_\alpha, w) - \lambda E_0(f_\alpha(v), x_\alpha, w)] \\ &= \sum_\alpha f_\alpha(v) E_0(\lambda, x_\alpha, w) = E_0(\lambda, v, w) \end{aligned}$$

because E_0 is a derivation in its first argument and linear in its second. This shows that $\Theta Q = E_0$ and concludes the proof.

Let $\mathfrak{D}$ be the set of all additive derivations of F, that is the set of all functions D from F to F which are additive and satisfy

$$D(\lambda\mu) = \lambda D(\mu) + \mu D(\lambda).$$

Let $\mathfrak{I}$ be the space of skew-symmetric bilinear forms on V. The map $\langle D, T\rangle \rightarrow E$, where $E(\lambda, v, w) = D(\lambda)T(v, w)$, is a bilinear map from $\mathfrak{D}\times\mathfrak{I}$ to $\mathcal{E}$ which induces an injective linear map from $\mathfrak{D}\otimes\mathfrak{I}$ to $\mathcal{E}$. (See [1] or [2] for the definition of the tensor product $\mathfrak{D}\otimes\mathfrak{I}$.) When V and hence $\mathfrak{I}$ is finite dimensional, it is easy to see that this map is also surjective. Hence we have the following corollary.

1.2 COROLLARY. *If V has finite dimension, $\mathfrak{Q}_0/\mathfrak{Q}$ is naturally isomorphic to* $\mathfrak{D}\times\mathfrak{I}$.

Let Q be a quasi-quadratic form and let S be the corresponding biadditive form. S will be homogeneous with respect to the scalar μ if and only if $E(\mu, v, w) = 0$ for all v and w. Since $\lambda\rightarrow E(\lambda, v, w)$ is a derivation of F for each v and w, S will certainly be homogeneous with respect to those scalars μ for which every derivation vanishes. Conversely, suppose that D is a derivation of F which does not vanish at μ and assume that V has dimension at least two. Define $E_0(\lambda, v, w) = D(\lambda)T(v, w)$, where T is some nontrivial skew-symmetric bilinear form on V. Any of the quasi-quadratic forms Q for which $\Theta Q = E_0$ will be associated with a symmetric biadditive form S which is not homogeneous with respect to the scalar μ.

Suppose that F has characteristic 0. Every derivation of F vanishes at every algebraic number, hence the form S associated with a quasi-quadratic form will be homogeneous with respect to algebraic numbers. On the other hand, if μ is transcendental, there is a derivation of F which does not vanish at μ (see [4], particularly Corollaries 1′ and 2′, pp. 124–5). Hence there are quasi-quadratic forms for which the associated biadditive form is not homogeneous with respect to μ.

Suppose that F has characteristic p. The set of all pth powers is a subfield F' of F and every derivation must vanish on all of F'. Moreover, if μ is in F but not F', there is a derivation of F which does not vanish at μ (see [4], particularly the remark on p. 126). Hence the biadditive form associated with a quasi-quadratic form is always homogeneous with respect to scalars in F' but need not be with respect to scalars not in F'.

2. Other identities for quadratic forms. Since the identities (2) and (4) fail to characterize quadratic forms, what other identity should we use? We can provide directly for the linearity of the biadditive function S by requiring the identity

$$Q(x + \lambda y) - Q(x - \lambda y) = \lambda(Q(x + y) - Q(x - y)), \tag{8}$$

in addition to the parallelogram law (2), since (8) becomes immediately $S(x, \lambda y) = \lambda S(x, y)$.

It turns out that (8) alone almost guarantees that Q is a quadratic form without requiring the parallelogram law. Both constant and linear functions satisfy the identity (8). We shall refer to a function as quadratic if it can be represented as the sum of a constant, a linear functional, and a quadratic form.

We shall see that (8) characterizes the quadratic functions except when F has three elements. (Note that every function from V to F satisfies (8) if F has only three elements.)

We begin with a simple extension of Corollary 0.2.

2.1 LEMMA. *Let V be a vector space over a field F of dimension at least two. Let Q be a function from V to F such that the restriction of Q to each two dimensional subspace of V is a quadratic function. Then Q is a quadratic function.*

Proof. Define

$$Q_1(x) = \tfrac{1}{2}(Q(x) + Q(-x) - 2Q(0)),$$
$$Q_2(x) = \tfrac{1}{2}(Q(x) - Q(-x)), \qquad \text{and}$$
$$Q_3(x) = Q(0).$$

A brief calculation shows that the restriction of Q_1 to each two dimensional subspace of V is a quadratic form. Hence, by corollary 0.2, Q_1 is a quadratic form. The restriction of Q_2 to each two dimensional subspace of V is a linear functional. Hence Q_2 is a linear functional. Since $Q=Q_1+Q_2+Q_3$, Q is a quadratic function.

2.2 THEOREM. *Let V be a vector space over a field F. Assume that F contains more than three elements and has characteristic different from 2. Every function Q from V to F satisfying the identity* (8) *is a quadratic function.*

Proof. In view of the lemma it is sufficient to prove this theorem when the dimension of V is zero, one, or two. When V has dimension zero, it is trivial.

Let Φ be the set of all functions ϕ from F to $\mathfrak{F}$ which satisfy the identity (8). If $\phi\in\Phi$ and $\lambda\in F$, then

$$\begin{aligned}\phi(2\lambda) &= \phi(0) + (\phi(\lambda + \lambda) - \phi(\lambda - \lambda)) \\ &= \phi(0) + \lambda(\phi(\lambda + 1) - \phi(\lambda - 1)) \\ &= \phi(0) + \lambda(\phi(1 + \lambda) - \phi(1 - \lambda)) + \lambda(\phi(0 + (1 - \lambda)) - \phi(0 - (1 - \lambda))) \\ &= \phi(0) + \lambda^2(\phi(2) - \phi(0)) + \lambda(1 - \lambda)(\phi(1) - \phi(-1)).\end{aligned}$$

This shows explicitly that Φ contains only quadratic functions. This proves the theorem for one dimensional vector spaces. It also shows that any member of Φ which vanishes at as many as three points is everywhere zero.

Now suppose that V is two dimensional. Let $\mathfrak{Q}_1$ be the set of all functions from V to F which satisfy (8). We know that $\mathfrak{Q}_1$ contains the six dimensional space of quadratic functions on V. It will therefore be sufficient to prove that $\mathfrak{Q}_1$ has dimension at most six over F.

Let L be a line in V, say $L=\{x+\lambda y \mid \lambda\in F\}$. Suppose that $Q\in\mathfrak{Q}_1$; then $\phi(\lambda)=Q(x+\lambda y)$ defines a function $\phi\in\Phi$. Hence, if Q vanishes at three points of L it vanishes at all points of L.

Let L_1, L_2, and L_3 be three lines of V in general position. Let x_1, x_2, and x_3

be their three points of intersection and let y_1, y_2, and y_3 be three other points on the lines L_1, L_2, and L_3, respectively. Suppose that $Q \in \mathfrak{Q}_1$ and Q vanishes at each of the six points x_1, x_2, x_3, y_1, y_2, and y_3. We shall prove that Q vanishes on all of V. This will prove that $\mathfrak{Q}_1$ has dimension at most six and thus finish the proof.

Since Q vanishes at three points of each of the lines L_1, L_2, and L_3, it vanishes at all points of $L_1 \cup L_2 \cup L_3$. Let z be any other point of V and assume F has at least seven elements. Among the at-least-eight lines through z, there must be one which does not contain x_1, x_2, or x_3 and is not parallel to L_1, L_2, or L_3. This line M meets L_1, L_2, and L_3 in three distinct points. Thus Q vanishes at three points of M. Therefore Q vanishes at z. We have proved that Q vanishes on all of V. The fields with two, three, and four elements have been excluded from the theorem. If F has five elements, it is readily checked that there are many lines which meet L_1, L_2, and L_3 in distinct points; then, repeating the argument, we see that Q vanishes on all of V.

The calculation involved in the proof of the lemma shows that we can distinguish the quadratic forms among the quadratic functions by requiring that

$$Q(-x) = Q(x) \tag{9}$$

and

$$Q(0) = 0, \tag{10}$$

both of which are implied by (4). Similarly, the linear functionals are singled out by the identity

$$Q(-x) = -Q(x). \tag{11}$$

With the aid of (9) we can write (8) in the form

$$Q(\lambda y + x) - Q(\lambda y - x) = \lambda(Q(y + x) - Q(y - x)). \tag{12}$$

This latter identity by itself implies (9) by taking $\lambda = 0$; hence (12) is equivalent to (8) and (9) together.

2.3 COROLLARY. *Let V be a vector space over a field F. Assume that F contains more than three elements and has characteristic different from 2. Then every function from V to F satisfying* (10) *and* (12) *is a quadratic form.*

The field of three elements must still be excluded because for this field (12) is equivalent to (9) which is insufficient to characterize quadratic forms.

Combining (8) and (11), we obtain the identity

$$Q(\lambda y + x) + Q(\lambda y - x) = \lambda(Q(y + x) + Q(y - x)). \tag{13}$$

Again setting $\lambda = 0$, we see that (13) is equivalent to (8) and (11). Therefore (13) characterizes linear functionals except over fields of characteristic two and the field of three elements.

References

1. N. Bourbaki, Eléments de Mathématique VII, Algèbre Chapitre 3, Algèbre Multilinéaire, Hermann et Cie., Paris, 1948.
2. C. Chevalley, Fundamental Concepts of Algebra, Academic Press, New York, 1956.
3. P. Jordan and J. von Neumann, On inner products in linear metric space, Annals of Math., 36 (1935) 719–723.
4. O. Zariski and P. Samuel, Commutative Algebra, Vol. 1, Van Nostrand, Princeton, 1958.

ON THE PROBABILITY THAT A RANDOM INTEGER HAS INITIAL DIGIT A

B. J. FLEHINGER, IBM, T. J. Watson Research Center, Yorktown Heights, New York

Several papers have been written discussing the distribution of first significant digits in tables of physical constants, [2, 3, 4, 6]. It has been well established empirically that in most such tables the proportion of numbers with first significant digit equal to or less than A is approximately $\log_{10}(A+1)$. That is, roughly three tenths of all known physical constants have one as the first significant digit and roughly seven tenths begin with digits less than five. Less than five hundredths begin with nine. Previous authors, such as Furry [3, 4] and Pinkham [6], have sought the explanation of this phenomenon by assuming that all physical constants are selected from a population with some underlying distribution and have shown that certain plausible assumptions about this distribution lead to the logarithmic law. Pinkham, for example, has shown that the only distribution of first significant digits which is invariant under scale change of the underlying distribution is $\log_{10}(A+1)$.

It occurred to this author that the smallest population which contains the set of significant figures of all possible physical constants, past, present, and future, must be the set of positive integers. The explanation for the logarithmic law should, therefore, lie in the properties of the set of integers as represented in a radix number system. Is there a meaningful way to answer the heuristic question, "What proportion of the positive integers have initial digit equal to or less than A?" or "What is the probability that an integer chosen at random has initial digit equal to or less than A?"

As a first step, consider those integers equal to or less than some finite number N and let $P_N^1(A)$ be the proportion of this subset which have initial digit equal to or less than A. If $\lim_{N\to\infty} P_N^1(A)$ existed, this limit would be a satisfactory answer to the questions posed above. However, as N increases $P_N^1(A)$ oscillates between $A/9$ (for $N=10^j$, integral j) and approximately $10\,A/9(A+1)$ (for $N=(A+1)\cdot 10^j$). The variation of $P_N^1(1)$ is plotted in Figure 1. Thus, $P_N^1(A)$ is a slowly divergent sequence, oscillating over successively longer periods. We are forced to seek the answer to our questions in a generalized or Banach limit of this sequence [1, 5].

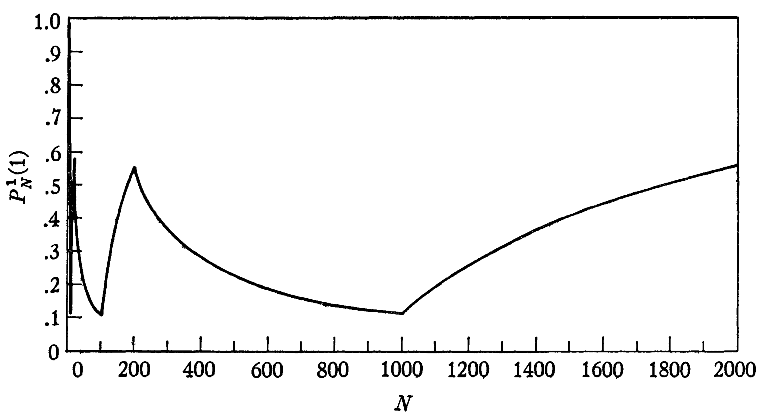


FIG. 1. Proportion of integers not exceeding N with initial digit equal to 1.

When we attempt to find a Cesaro limit of the sequence, i.e. form the cumulative averages

$$P_N^2(A) = \frac{1}{N} \sum_{M=1}^{N} P_M^1(A)$$

we again generate a divergent sequence, but one which oscillates between narrower limits. As we take successive cumulative averages, (Hölder sums)

$$P_N^k(A) = \frac{1}{N} \sum_{M=1}^{N} P_M^{k-1}(A),$$

we generate divergent sequences for all finite k, but, as k increases, the limits within which the sequences oscillate for large N become closer together. This behavior is illustrated in Figure 2. It will be proved in the following section that

$$\lim_{k\to\infty} \liminf_{N\to\infty} P_N^k(A) = \lim_{k\to\infty} \limsup_{N\to\infty} P_N^k(A) = \log_{10}(A+1).$$

Thus, we have found a regular limiting process which leads to a probability measure on the set of integers with initial digit equal to or less than A. This measure agrees with the logarithmic law.

This discussion is limited to the base ten number system, but it can obviously be generalized to the case of any radix.

Mathematical development. Let $S(A)$ be the set of positive integers with initial digit equal to or less than A, and let $P_N^0(A)$ be the characteristic function of $S(A)$, i.e.,

$$P_N^0(A) = \begin{cases} 1 & N \in S(A) \\ 0 & N \notin S(A). \end{cases} \tag{1}$$

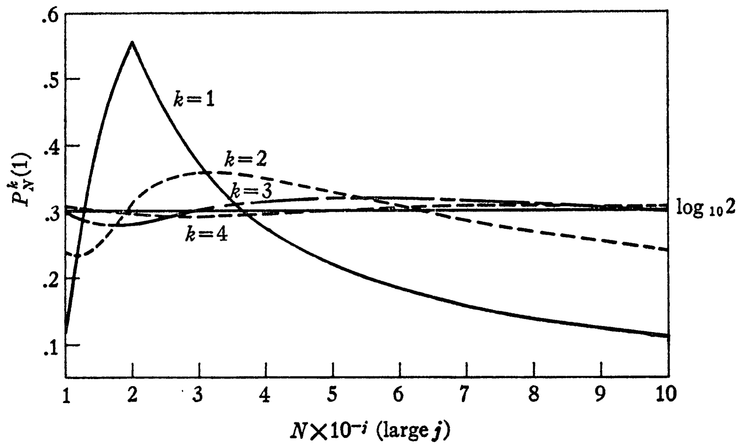


FIG. 2. Variation of $P_N^k(1)$ as N varies from 10^j to 10^{j+1} (large j).

Let

$$P_N^k(A) = \frac{1}{N}\sum_{M=1}^{N} P_M^{k-1}(A). \tag{2}$$

Clearly, $P_N^1(A)$ is the proportion of positive integers bounded by N which are members of $S(A)$.

It will now be proved that

$$\lim_{k\to\infty}\liminf_{N\to\infty} P_N^k(A) = \lim_{k\to\infty}\limsup_{N\to\infty} P_N^k(A) = \log_{10}(A+1). \tag{3}$$

Note that

$$P_N^1(A) = \begin{cases} 1 - \dfrac{(9-A)(10^j-1)}{9N} & 10^j \leqq N < (A+1)\cdot 10^j (N \in S(A)) \\ \dfrac{A(10^{j+1}-1)}{9N} & (A+1)\cdot 10^j \leqq N < 10^{j+1} (N \notin S(A)). \end{cases} \tag{4}$$

Although $P_N^1(A)$ does not approach a limit as N increases, $P^1_{\alpha\cdot 10^j}(A)$ does converge to a limit for any α between 1 and 10 as j goes to infinity and

$$Q^1(\alpha, A) = \lim_{j\to\infty} P^1_{\alpha\cdot 10^j}(A) = \begin{cases} 1 - \dfrac{9-A}{9\alpha} & 1 \leqq \alpha < A+1 \\ \dfrac{10A}{9\alpha} & A+1 \leqq \alpha < 10. \end{cases} \tag{5}$$

Furthermore, we observe that, for all k, $\lim_{j\to\infty} P^k_{\alpha\cdot 10^j}(A)$ exists, and we set this limit equal to $Q^k(\alpha, A)$. Then

$$
\begin{aligned}
\liminf_N P_N^k(A) &= \min_{1\leqq\alpha<10} Q^k(\alpha, A) \\
\limsup_N P_N^k(A) &= \max_{1\leqq\alpha<10} Q^k(\alpha, A).
\end{aligned}
\tag{6}
$$

Therefore, we can prove (3) by showing that $\lim_{k\to\infty} Q^k(\alpha, A) = \log_{10}(A+1)$, $1\leqq\alpha<10$.

We now derive an expression for $Q^k(\alpha, A)$ for $k>1$.

$$
\begin{aligned}
Q^k(\alpha, A) &= \lim_{j\to\infty} P^k_{\alpha\cdot 10^j}(A) = \lim_{j\to\infty} \frac{1}{\alpha\cdot 10^j} \sum_{M=1}^{\alpha\cdot 10^j} P_M^{k-1}(A) \\
&= \lim_{j\to\infty} \frac{1}{\alpha\cdot 10^j}\left[\sum_{r=0}^{j-1} \sum_{M=10^r}^{10^{r+1}-1} P_M^{k-1}(A) + \sum_{M=10^j}^{\alpha\cdot 10^j} P_M^{k-1}(A)\right] \\
&= \frac{1}{\alpha}\left[\sum_{s=1}^{\infty} 10^{-s} \int_1^{10} Q^{k-1}(\beta, A)\,d\beta + \int_1^{\alpha} Q^{k-1}(\beta, A)\,d\beta\right]\cdot
\end{aligned}
\tag{7}
$$

Thus,

$$
Q^k(\alpha, A) = \frac{1}{\alpha}\left[\frac{1}{9}\int_1^{10} Q^{k-1}(\beta, A)\,d\beta + \int_1^{\alpha} Q^{k-1}(\beta, A)\,d\beta\right]\cdot
\tag{8}
$$

Starting with (5), the successive application of this relationship yields expressions for $Q^k(\alpha, A)$ for all k. We shall prove by induction that

$$
Q^k(\alpha, A) = \begin{cases}
1 - \dfrac{1}{\alpha}\displaystyle\sum_{j=0}^{k-1} \frac{c_{k-j}(A)(\log\alpha)^j}{j!}, & 1 \leqq \alpha < A+1, \\
\dfrac{1}{\alpha}\displaystyle\sum_{j=0}^{k-1} \frac{d_{k-j}(A)(\log\alpha)^j}{j!}, & A+1 \leqq \alpha < 10,
\end{cases}
\tag{9}
$$

where the generating functions associated with $\{c_j(A)\}$ and $\{d_j(A)\}$ are

$$
\begin{aligned}
C(Z, A) &= \sum_{j=1}^{\infty} c_j(A) Z^j = \frac{Z}{1-Z}\,\frac{10^{1-Z} - (A+1)^{1-Z}}{10^{1-Z}-1} \\
D(Z, A) &= \sum_{j=1}^{\infty} d_j(A) Z^j = \frac{Z}{1-Z}\,\frac{10^{1-Z}[(A+1)^{1-Z}-1]}{10^{1-Z}-1}.
\end{aligned}
\tag{10}
$$

From (5), we observe that $Q^1(\alpha, A)$ satisfies (9) with

$$
c_1(A) = \frac{9-A}{9} = \lim_{Z\to 0} Z^{-1}C(Z, A),\quad d_1(A) = \frac{10A}{9} = \lim_{Z\to 0} Z^{-1}D(Z, A).
$$

Now assume (9) for $Q^{k-1}(\alpha, A)$. Then, substituting this expression in (8), we verify (9) for $Q^k(\alpha, A)$ and derive the following expressions for $c_k(A)$ and $d_k(A)$:

$$\begin{aligned} c_k(A) &= \frac{9-A}{9} + \frac{1}{9}\sum_{j=1}^{k-1} \frac{[c_{k-j}(A) + d_{k-j}(A)][\log(A+1)]^j}{j!} \\ &\quad - \frac{1}{9}\sum_{j=1}^{k-1} \frac{d_{k-j}(A)(\log 10)^j}{j!} \\ d_k(A) &= \frac{10A}{9} - \frac{10}{9}\sum_{j=1}^{k-1} \frac{[c_{k-j}(A) + d_{k-j}(A)][\log(A+1)]^j}{j!} \\ &\quad + \frac{1}{9}\sum_{j=1}^{k-1} \frac{d_{k-j}(A)(\log 10)^j}{j!}. \end{aligned} \tag{11}$$

Thus the generating functions satisfy

$$\begin{aligned} C(Z, A) &= \frac{9-A}{9}\,\frac{Z}{1-Z} + \frac{1}{9}[C(Z, A) + D(Z, A)][(A+1)^Z - 1] \\ &\quad - \frac{1}{9} D(Z, A)[10^Z - 1] \\ D(Z, A) &= \frac{10A}{9}\,\frac{Z}{1-Z} - \frac{10}{9}[C(Z, A) + D(Z, A)][(A+1)^Z - 1] \\ &\quad + \frac{1}{9} D(Z, A)[10^Z - 1], \end{aligned} \tag{12}$$

which may be solved to yield (10).

We now consider the generating function associated with the set of functions $\{Q^k(\alpha, A)\}$. $\phi(Z, \alpha, A) = \sum_{k=1}^{\infty} Q^k(\alpha, A)Z^k$.

From (9),

$$\phi(Z, \alpha, A) = \begin{cases} \dfrac{Z}{1-Z} - \dfrac{C(Z, A)}{\alpha^{1-Z}} & 1 \leqq \alpha < A+1 \\[2ex] \dfrac{D(Z, A)}{\alpha^{1-Z}} & A+1 \leqq \alpha < 10. \end{cases} \tag{13}$$

Upon substituting (10) we obtain

$$\phi(Z, \alpha, A) = \begin{cases} \dfrac{Z}{1-Z}\left[1 - \dfrac{10^{1-Z} - (A+1)^{1-Z}}{\alpha^{1-Z}(10^{1-Z} - 1)}\right] & 1 \leqq \alpha < A+1 \\[2ex] \dfrac{Z}{1-Z}\,\dfrac{10^{1-Z}[(A+1)^{1-Z} - 1]}{\alpha^{1-Z}[10^{1-Z} - 1]} & A+1 \leqq \alpha < 10. \end{cases} \tag{14}$$

Now

$$Q^k(\alpha, A) = \frac{1}{2\pi i}\int_\Gamma \frac{1}{Z^{k+1}}\phi(Z, \alpha, A)dZ, \tag{15}$$

where Γ is any contour containing the origin and not containing any poles of $\phi(Z, \alpha, A)$. For $1 \leqq \alpha < 10$, $\phi(Z, \alpha, A)$ has simple poles only at the points

$$1 + \frac{2\pi n i}{\log 10}, \quad n = 0, \pm 1, \pm 2, \cdots.$$

We let $\Gamma = \Gamma' - \Gamma''$ where Γ' is a circle centered at the origin with radius R, where $1 < R < 2$, and Γ'' is a circle centered at 1 with radius $R-1$. Then

$$\begin{aligned} Q^k(\alpha) &= \frac{1}{2\pi i}\int_{\Gamma'} \frac{1}{Z^{k+1}}\phi(Z, \alpha, A)dZ - \frac{1}{2\pi i}\int_{\Gamma''} \frac{1}{Z^{k+1}}\phi(Z, \alpha, A)dZ \\ &= O(R^{-(k-1)}) - \lim_{Z\to 1}\frac{Z-1}{Z^{k+1}}\phi(Z, \alpha, A). \end{aligned} \tag{16}$$

Using L'Hôpital's Rule to take these limits, we find that, for $1 \leqq \alpha < 10$, $Q^k(\alpha, A) = (\log (A+1)/\log 10) + O(R^{-(k-1)})$, so that

$$\lim_{k\to\infty} Q^k(\alpha, A) = \log_{10}(A+1), \tag{17}$$

which completes the proof.

References

1. R. L. Adler and A. G. Konheim, Solution of problem 4999, Advanced problems and solutions, this MONTHLY, 70 (1963) 218–219.

2. Frank Benford, The law of anomalous numbers, Proc. Amer. Philos. Soc., 78 (1938) 551–572.

3. W. H. Furry and H. Hurwitz, Distribution of numbers and distribution of significant figures, Nature, 155 (1945) 52–53.

4. S. A. Goudsmit and W. H. Furry, Significant figures of numbers in statistical tables, Nature, 154 (1944) 800–801.

5. G. H. Hardy, Divergent Series, Oxford University Press, London, 1956, Chapter V.

6. Roger S. Pinkham, On the distribution of first significant digits, Ann. Math. Stat., 32 (1961) 1223–1230.

Mathematical Swifties

"An interpretation of that axiom system exists," remarked Tom with satisfaction.

"$L = \int_\alpha^\beta \sqrt{1+(y')^2}\,dx$," said Tom at length.

CONSTRUCTION OF A CHARACTERISTIC BASIS FOR A MATRIX

C. F. MOPPERT, Melbourne University

1. Introduction. Let M be the set of $n\times n$ matrices with complex coefficients, let e be the unit matrix and let $\epsilon_{\nu\mu}$ $(\nu, \mu=1, \cdots, n)$ be the matrix with 1 in the intersection of row ν and column μ. The set M is an algebra over the field of complex numbers and the matrices $\epsilon_{\nu\mu}$ form a basis of it. This basis we call the *ordinary basis* of the algebra M. Any set $\epsilon^*_{\nu\mu}$ $(\nu, \mu=1, \cdots, n)$ of matrices of rank 1 of M satisfying the multiplication table $\epsilon^*_{\nu\mu}\epsilon^*_{\rho\sigma}=\delta_{\mu\rho}\epsilon^*_{\nu\sigma}$ (Kronecker δ) we call a *normal basis* of M.

If a is any matrix of M in the Jordan normal form then we say that the ordinary basis of M is *characteristic for the matrix a*.

Again, let $\epsilon^*_{\nu\mu}$ be a normal basis and $\epsilon_{\nu\mu}$ be the ordinary basis of M. Let T be an automorphism of M such that $T\epsilon^*_{\nu\mu}=\epsilon_{\nu\mu}$ holds for $\nu, \mu=1, \cdots, n$. Let a be an element of M such that the ordinary basis is characteristic for Ta. In this case, we call *the basis $\epsilon^*_{\nu\mu}$ characteristic for the matrix a.*

In a total matrix algebra, any automorphism is an inner automorphism [**1**]. Accordingly, if the basis $\epsilon^*_{\nu\mu}$ is characteristic for a given matrix a, then there is a regular $t\in M$ such that the matrix tat^{-1} appears in the Jordan normal form.

We shall show in this paper how to construct a basis of M which is characteristic for a given matrix a, naturally without presupposing the possibility of the representation of a in the Jordan normal form. To do so, we must first define the characteristic basis in an invariant manner.

If $\lambda_1, \cdots, \lambda_r$ are the distinct eigenvalues of a, then it is well known that there are idempotent elements $e_{10}, \cdots, e_{r0}$ and nilpotent elements $e_{11}, \cdots, e_{r1}$ such that the relations

$$
\begin{aligned}
a &= \sum_{i=1}^{r} (\lambda_i e_{i0} + e_{i1}) \\
e &= \sum_{i=1}^{r} e_{i0} \qquad (1)\\
\left.\begin{aligned} e_{i0}e_{k0} &= \delta_{ik}e_{i0} \\ e_{i0}e_{i1}e_{i0} &= e_{i1} \end{aligned}\right\} & \ (i, k = 1, \cdots, r)
\end{aligned}
$$

hold. All these elements can be represented as polynomials in a with complex coefficients and can be calculated by using the minimum polynomial of a only. Our construction of the basis characteristic for a will start from these elements.

It might be appropriate to recall how the representation (1) can be found. Let $m(x)=(x-\lambda_1)^{\mu_1}, \cdots, (x-\lambda_r)^{\mu_r}$ be the minimum polynomial of a. We develop $(m(x))^{-1}$ in partial fractions:

$$\frac{1}{m(x)}=\frac{\alpha_1}{x-\lambda_1}+\cdots+\frac{\alpha_{\mu_1}}{(x-\lambda_1)^{\mu_1}}+\frac{\beta_1}{(x-\lambda_2)}+\cdots.$$

We then put

$$e_{1i}(x) = (x - \lambda_1)^i m(x) \left(\frac{\alpha_1}{x - \lambda_1} + \cdots + \frac{\alpha_{\mu_1}}{(x - \lambda_1)^{\mu_1}} \right) \qquad (i = 0, \cdots, \mu_1 - 1).$$

From this we get the matrices e_{1i} $(i=0, \cdots, \mu_1-1)$ by replacing x by the matrix a and 1 by the matrix e. The remaining $e_{\nu\mu}$ we get in the corresponding manner.

For our purposes it is essential that the representation (1) gives a matrix a as a linear combination of idempotent and nilpotent matrices, all these matrices being polynomials in the matrix a [2].

1. Notations and definitions. In the following, M will denote a total matrix algebra with unit e over the field of complex numbers. The elements of M will be denoted by small latin letters or else by ϵ, ϵ_ν, $\epsilon_{\nu\mu}$. Any element denoted by ϵ (with or without a subscript) will be of rank 1. Elements denoted by ϵ, ϵ_ν or by $\epsilon_{\nu\nu}$ will furthermore be idempotent elements. Complex numbers will be denoted by small Greek letters (no ϵ will appear in this context). Small latin letters will be used also for enumerating purposes.

An element $a \in M$ is called *nilpotent of order k* if $a^k=0$ and $a^{k-1} \neq 0$. If n^2 is the order of M and if a is nilpotent of order n then we say that a is *maximal nilpotent.*

If e_1 and e_2 are idempotent elements of M with $e_1e_2=e_2e_1=0$ then we say that e_1 and e_2 are *orthogonal.* If e' is idempotent and if the elements $e_1, \cdots, e_k$ are pairwise orthogonal idempotent with $e'=e_1+\cdots+e_k$ then we call this an *orthogonal decomposition of* e'. If furthermore all the elements $e_1, \cdots, e_k$ are of rank 1 then we call the orthogonal representation mentioned above *primitive orthogonal.*

If $a \in M$ is a nilpotent element, then we call an *orthogonal decomposition* $e=e_1+\cdots+e_k$ *of e characteristic for a* if the ranks of the e_i's are monotonously decreasing, if $a=e_1ae_1+\cdots+e_kae_k$, and if each e_iae_i $(i=1, \cdots, k)$ is maximal nilpotent in e_iMe_i.

Let e' and e'' be orthogonal idempotent elements of M and let r', r'' be their ranks. If then $\epsilon'_{\nu\mu}$ $(\nu, \mu=1, \cdots, r')$ is a normal basis of $e'Me'$ and $\epsilon''_{\nu\mu}$ $(\nu, \mu=1, \cdots, r'')$ is a normal basis of $e''Me''$, then we call a normal basis $\epsilon_{\nu\mu}$ $(\nu, \mu=1, \cdots, r'+r'')$ of $(e'+e'')M(e'+e'')$ an *extension of the basis* $\epsilon'_{\nu\mu}$ *via* $\epsilon''_{\nu\mu}$ if $\epsilon_{\nu\mu}=\epsilon'_{\nu\mu}$ holds for $\nu, \mu=1, \cdots, r'$ and $\epsilon_{\nu\mu}=\epsilon''_{\nu-r',\mu-r'}$ holds for $\nu, \mu=r'+1, \cdots, r'+r''$.

The definition of a *basis characteristic for a given matrix* will be given in four steps.

a) If the matrix a is maximal nilpotent in M where M has order n^2 then we call a normal basis $\epsilon_{\nu\mu}$ $(\nu, \mu=1, \cdots, n)$ of M characteristic for a provided $a=\epsilon_{12}+\epsilon_{23}+\cdots+\epsilon_{n-1,n}$.

b) If the matrix a is nilpotent in M and if the decomposition of e, namely $e=e_1+\cdots+e_k$ is for a characteristic, then we call a normal basis of M char-

acteristic for a provided it satisfies the following conditions: if $\epsilon'_{\nu\mu}$ is a characteristic basis of $(e_1+\cdots+e_i)M(e_1+\cdots+e_i)$ for $(e_1+\cdots+e_i)a(e_1+\cdots+e_i)$ and if $\epsilon'_{\nu\mu}$ is a characteristic basis of $e_{i+1}Me_{i+1}$ for $e_{i+1}ae_{i+1}$, then for every $i=1, \cdots, k-1$ the basis $\epsilon_{\nu\mu}$ of $(e_1+\cdots+e_{i+1})M(e_1+\cdots+e_{i+1})$ is an extension of the basis $\epsilon'_{\nu\mu}$ via $\epsilon''_{\nu\mu}$.

c) If a has exactly one eigenvalue λ, then we call a normal basis of M characteristic for a provided it is a characteristic basis for the matrix $a-\lambda e$.

d) If $\lambda_1, \cdots, \lambda_r$ are the distinct eigenvalues of the matrix a and if a is given by the representation (1) then we call a normal basis of M characteristic for a provided it is defined in a recursive manner in the sense of b), the elements $e_1, \cdots, e_k$ occurring there being replaced by $e_{10}, \cdots, e_{r0}$.

2. Subsidiary lemmas. We shall denote by R right ideals and by L left ideals in M. The set of all linear combinations with complex coefficients of elements $a, b, c, \cdots \in M$ we shall denote by $\bigvee(a, b, c, \cdots)$.

LEMMA 1. *If R, L are nonzero simple ideals in M then $R\cap L\neq 0$. Furthermore, any two elements of $R\cap L$ are linearly dependent.*

Proof. Let n^2 be the order of M and let $\epsilon_{\nu\mu}$ be a normal basis of M. To a simple R is associated a set $\alpha_1, \cdots, \alpha_n$ in such a way that any $x\in R$ can be represented in the form $\sum\alpha_\nu\xi_\mu\epsilon_{\nu\mu}$. Similarly, to any simple L is associated a set $\beta_1, \cdots, \beta_n$ in such a way that $x=\sum\beta_\mu\eta_\nu\epsilon_{\nu\mu}$ holds for any $x\in L$. Accordingly, $u=\sum\alpha_\nu\beta_\mu\epsilon_{\nu\mu}$ $\in R\cap L$ and we have $u\neq 0$ provided neither all the α's nor all the β's vanish.

An element $x=\sum\alpha_\nu\xi_\mu\epsilon_{\nu\mu}$ generates the left ideal Mx with elements of the form $y=\sum\xi_\mu\eta_\nu\epsilon_{\nu\mu}$. Accordingly, $Mx=L$ if and only if $\xi_\mu=\xi\beta_\mu$. It follows that $x\in R\cap L$ if and only if $x=\xi u$. This completes the proof of Lemma 1.

COROLLARY 1. *If u is of rank 1 then $u^2=\xi u$.*

Proof. Both the ideals $R=uM$ and $L=Mu$ are simple. Both the elements u and u^2 are in $R\cap L$.

COROLLARY 2. *To any ϵ_1, ϵ_2 there is a $u\neq 0$ such that $u\in\epsilon_1M\epsilon_2$.*

Proof. The ideals $R=\epsilon_1M$ and $L=M\epsilon_2$ are both simple. Accordingly, there is a $u\neq 0$, $u\in R\cap L$, $u=\epsilon_1u=u\epsilon_2\in\epsilon_1M\epsilon_2$.

LEMMA 2. *If a, b, ϵ are elements such that $a\epsilon\neq 0$ and $\epsilon b\neq 0$ then $a\epsilon b\neq 0$.*

Proof. From $a\epsilon\in M\epsilon$ it follows that $Ma\epsilon=M\epsilon$ and that $\epsilon\in Ma\epsilon$. Accordingly, there is an s such that $\epsilon=sa\epsilon$. Similarly, there is a t such that $\epsilon=\epsilon bt$. It follows $\epsilon=\epsilon^2=sa\epsilon bt$ and from this the lemma ensues.

LEMMA 3. *If $\epsilon_1, \cdots, \epsilon_k$ are pairwise orthogonal elements and if $u_\nu\neq 0$, $u_\nu=\epsilon_\nu u_\nu\epsilon_{\nu+1}$ $(\nu=1, \cdots, k-1)$, then there is a normal basis $\epsilon_{\nu\mu}$ $(\nu, \mu=1, \cdots, k)$ of $(\epsilon_1+\cdots+\epsilon_k)M(\epsilon_1+\cdots+\epsilon_k)$ such that $\epsilon_{\nu\nu}=\epsilon_\nu$ and $\epsilon_{\nu,\nu+1}=u_\nu$.*

Proof. According to Lemma 2, $u_\nu u_{\nu+1} \cdots u_{\mu-1} \neq 0$ for $\nu = 1, \cdots, k-1$; $\mu = 2, \cdots, k$; $\mu > \nu$. We define $\epsilon_{\nu\mu} = u_\nu \cdots u_{\mu-1}$ for $\mu > \nu$ and $\epsilon_{\nu\nu} = \epsilon_\nu$. It follows that $\epsilon_{\nu\mu}\epsilon_{\rho\sigma} = \delta_{\mu\rho}\epsilon_{\nu\sigma}$ holds for $\mu \geqq \nu$, $\sigma \geqq \rho$. For $\nu = 1, \cdots, k-1$ we choose $v_\nu' \neq 0$, $v_\nu' \in \epsilon_{\nu+1} M \epsilon_\nu$. According to Lemma 2 we have then $u_\nu v_\nu' \neq 0$ and as $u_\nu v_\nu' \in \epsilon_\nu M \cap M\epsilon_\nu$, it follows that $u_\nu v_\nu' = \gamma_\nu \epsilon_\nu$ with $\gamma_\nu \neq 0$. Putting $v_\nu = (1/\gamma_\nu)v_\nu'$ it follows that $u_\nu v_\nu = \epsilon_\nu$. Then $v_\nu u_\nu v_\nu u_\nu = v_\nu \epsilon_\nu u_\nu = v_\nu u_\nu$; thus $v_\nu u_\nu = \epsilon_{\nu+1}$, as $v_\nu u_\nu \in \epsilon_{\nu+1} M \cap M\epsilon_{\nu+1}$. We now define $\epsilon_{\nu\mu} = v_{\nu-1} v_{\nu-2} \cdots v_\mu$ for $\nu > \mu$. This is a basis satisfying the conditions.

COROLLARY. *Let e', e'' be orthogonal idempotent elements of M, let $\epsilon_{\nu\mu}'$ $(\nu, \mu = 1, \cdots, r')$ be a normal basis of $e'Me'$ and $\epsilon_{\nu\mu}''$ $(\nu, \mu = 1, \cdots, r'')$ be a normal basis of $e''Me''$. Then there is a normal basis $\epsilon_{\nu\mu}(\nu, \mu = 1, \cdots, r'+r'')$ of $(e'+e'')M(e'+e'')$ which is an extension of the basis $\epsilon_{\nu\mu}'$ via $\epsilon_{\nu\mu}''$.*

Proof. Let u be a nonvanishing element of $\epsilon_{r',r'}'M \cap M\epsilon_{11}''$. We put $\epsilon_{\nu\nu} = \epsilon_{\nu\nu}'$ for $\nu = 1, \cdots, r'$ and $\epsilon_{\nu\nu} = \epsilon_{\nu-r',\nu-r'}''$ for $\nu = r'+1, \cdots, r'+r''$. Furthermore, we put $\epsilon_{\nu,\nu+1} = \epsilon_{\nu,\nu+1}'$ for $\nu = 1, \cdots, r'-1$; $\epsilon_{r',r'+1} = u$ and we put $\epsilon_{\nu,\nu+1} = \epsilon_{\nu-r',\nu+1-r'}''$ for $\nu = r'+1, \cdots, r'+r''-1$. Applying Lemma 3 we get the Corollary.

LEMMA 4. *Let n^2 be the order of M, let R, L be simple ideals and let u be a nonvanishing element of $R \cap L$. Furthermore, let $a_1, \cdots, a_{n-1}$, u be linearly independent elements of R. There is then an $\epsilon \in L$ such that $a_\nu \epsilon = 0$ for $\nu = 1, \cdots, n-1$.*

Proof. In $u^2 = \xi u$ we distinguish the two cases $\xi \neq 0$ and $\xi = 0$.

a) $\xi \neq 0$. Without restriction of generality we may assume that $u = \epsilon_1$. There is then $R = \epsilon_1 M$, $L = M\epsilon_1$. Let $e = \epsilon_1 + \cdots + \epsilon_n$ be a primitive orthogonal decomposition of e. We choose u_ν $(\nu = 1, \cdots, n)$ in such a way that $u_\nu \neq 0$ and that $u_\nu = \epsilon_1 u_\nu \epsilon_\nu$; in particular $u_1 = \epsilon_1$. If then the elements $a_2, \cdots, a_n$ are in R then we have $a_\nu = \alpha_{\nu 1} u_1 + \cdots + \alpha_{\nu n} u_n$ for $\nu = 2, \cdots, n$, furthermore are the elements $\epsilon_1, a_2, \cdots, a_n$ linearly independent if and only if

$$\begin{vmatrix} 1 & 0 & \cdots 0 \\ \alpha_{21} & \alpha_{22} & \cdots \alpha_{2n} \\ \cdot & \cdot & \cdots \cdot \\ \alpha_{n1} & \alpha_{n2} & \cdots \alpha_{nn} \end{vmatrix} \neq 0,$$

the elements $u_1, \cdots, u_n$ being linearly independent. We now choose nonvanishing elements v_ν', $v_\nu' \in \epsilon_\nu M \cap M\epsilon_1$ $(\nu = 2, \cdots, n)$. These are contained in L and $u_\nu v_\nu' \neq 0$ for $\nu = 2, \cdots, n$. Applying the argument used above we see that $u_\nu v_\nu' = \gamma_\nu \epsilon_1$ with $\gamma_\nu \neq 0$. Again, we put $v_\nu = (1/\gamma_\nu)v_\nu'$ for $\nu = 2, \cdots, n$ and $v_1 = \epsilon_1$. Accordingly we have then $u_\nu v_\mu = \delta_{\nu\mu}\epsilon_1$ for $\nu, \mu = 1, \cdots, n$.

The elements $v_1, \cdots, v_n$ of L are linearly independent and $v_\nu v_\mu = 0$ for $\mu \neq 1$ while $v_\nu v_1 = v_\nu$. We put $\epsilon = \xi_1 v_1 + \cdots + \xi_n v_n$. Then $a_\nu \epsilon = (\xi_1 \alpha_{\nu 1} + \cdots + \xi_n \alpha_{\nu n})\epsilon_1$, and ϵ satisfies the conditions of Lemma 4 provided the quantities $\xi_1, \cdots, \xi_n$ satisfy the system

$$\begin{aligned} \xi_1 \qquad\qquad\qquad &= 1 \\ \xi_1\alpha_{21} + \cdots + \xi_n\alpha_{2n} &= 0 \\ \cdots\cdots\cdots\cdots & \\ \xi_1\alpha_{n1} + \cdots + \xi_n\alpha_{nn} &= 0. \end{aligned}$$

This system has a solution because its determinant does not vanish. Finally we denote a_n by a_1 and have thus proved Lemma 4 in this case.

b) $\xi=0$. We choose $\epsilon_1 \in R$, $\epsilon_n' \in L$. Accordingly, $R=\epsilon_1 M$, $L=M\epsilon_n'$. There are elements x, y' such that $\epsilon_1 = ux$, $\epsilon_n' = y'u$. As we have $u^2=0$ it follows that $u\epsilon_1=\epsilon_n'\epsilon_1=0$. Furthermore, $\epsilon_1\epsilon_n' = \tau u$. We put $\epsilon_n=\epsilon_n'-\tau u$ and have now $\epsilon_1\epsilon_n=\epsilon_n\epsilon_1=0$.

Let $e=\epsilon_1+\cdots+\epsilon_{n-1}+\epsilon_n$ be a primitive orthogonal decomposition of e. Then there are nonvanishing elements u_ν $(\nu=1,\cdots,n)$ in such a way that $u_\nu=\epsilon_1 u_\nu \epsilon_\nu$ holds. In particular, we choose $u_n=u$. These are then linearly independent elements of R, and the elements a_ν can be represented in the form $a_\nu=\alpha_{\nu 1}u_1+\cdots+\alpha_{\nu n}u_n$. By assumption, the elements $a_1,\cdots,a_{n-1}$, u are linearly independent, i.e.,

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \cdot & \cdots & \cdot \\ \alpha_{n-1,1} & \cdots & \alpha_{n-1,n} \\ 0 & \cdots & 1 \end{vmatrix} \neq 0.$$

Again, we choose elements $v_\nu' \neq 0$ $(\nu=1,\cdots,n)$ in such a way that $\epsilon_\nu v_\nu' \epsilon_n = v_\nu'$ holds. These elements are contained in L and we have $u_\nu v_\mu'=0$ for $\nu\neq\mu$. Since $u_\nu v_\nu'\neq 0$, $u_\nu v_\nu'=\gamma_\nu u$ with $\gamma_\nu\neq 0$. Putting $v_\nu=(1/\gamma_\nu)v_\nu'$ we have $u_\nu v_\mu=\delta_{\nu\mu}u$ for ν, $\mu=1,\cdots,n$. In particular we have $u_n v_n=u=u_n$; $v_n=\epsilon_n$. The elements $v_1,\cdots,v_n$ of L are linearly independent and we put $\epsilon=\xi_1 v_1+\cdots+\xi_n v_n$. Then $a_\nu\epsilon=(\xi_1\alpha_{\nu 1}+\cdots+\xi_n\alpha_{\nu n})u$ for $\nu=1,\cdots,n-1$. The element ϵ satisfies the conditions of Lemma 4 provided

$$\begin{aligned} \xi_1\alpha_{11} + \cdots + \xi_n\alpha_{1n} &= 0 \\ \cdots\cdots\cdots\cdots & \\ \xi_1\alpha_{n-1,1} + \cdots + \xi_n\alpha_{n-1,n} &= 0 \\ \xi_n &= 1. \end{aligned}$$

This system can be solved, for its determinant does not vanish.

LEMMA 5. *If a is any element of M and if ϵ is contained in aM, then* $M\epsilon a \cap M(a-\epsilon a)=0$.

Proof. There is an a_1 such that $\epsilon=aa_1$ holds, thus $aa_1=aa_1aa_1$. Let x be a nonvanishing element in $M\epsilon a$. There is then an element x_1 such that $x=x_1\epsilon a$ $=x_1aa_1a$. It follows that $x_1aa_1=x_1aa_1aa_1=xa_1\neq 0$. On the other hand, if x is

contained in $M(a-\epsilon a)$ then there is an element x_2 such that $x=x_2(a-\epsilon a)=x_2(a-aa_1a)$. It follows that $xa_1=0$. Accordingly, $x\neq 0$, $x\in M\epsilon a\cap M(a-\epsilon a)$ is impossible.

COROLLARY. *Let a be nilpotent of order k, $k>1$ and let ϵ be contained in $a^{k-1}M$. Put $R=\epsilon M$ and let u be a non-vanishing element of $R\cap M\epsilon a^{k-1}$. If the elements $u_1, \cdots, u_p$ are linearly independent and if $R\cap M(a^{k-1}-\epsilon a^{k-1})=\mathsf{V}(u_1, \cdots, u_p)$, then the elements $\epsilon, \epsilon a, \cdots, \epsilon a^{k-2}, u_1, \cdots, u_p, u$ are linearly independent.*

Proof. Obviously, $u_1a= \cdots =u_pa=ua=0$. Suppose that

$$\xi_0\epsilon + \xi_1\epsilon a + \cdots + \xi_{k-2}\epsilon a^{k-2} + \xi_{k-1}u_1 + \cdots + \xi_{k+p-2}u_p + \xi_{k+p+1}u = 0.$$

Multiplication by a^{k-1} from the right gives $\xi_0\epsilon a^{k-1}=0$, thus $\xi_0=0$. Similarly we get $\xi_1= \cdots =\xi_{k-2}=0$. According to Lemma 5 we have $M\epsilon a^{k-1}\cap M(a^{k-1}-\epsilon a^{k-1})=0$, i.e., u is not contained in $\mathsf{V}(u_1, \cdots, u_p)$; thus $\xi_{k-1}= \cdots =\xi_{k-p+1}=0$.

3. Construction of a characteristic basis.

THEOREM 1. *If a is nilpotent of order k then there is an idempotent element e_1 of rank k such that $e_1a=ae_1$, this element being nilpotent of order k. There is furthermore a normal basis of e_1Me_1 which is characteristic for e_1ae_1.*

Proof. We are going to construct e_1 together with the characteristic base in three steps (the trivial case $k=1$ we take for granted).

a) We choose ϵ_1 in $a^{k-1}M$. There is then an element x such that $\epsilon_1=a^{k-1}x=\epsilon_1a^{k-1}x$. We put $a'=\epsilon_1a^{k-1}$; obviously $a'\neq 0$. Then we put $a''=a^{k-1}-a'=(e-\epsilon_1)a^{k-1}$. It follows that $a'a=a''a=0$, and from $a\epsilon_1=0$ it follows that $aa'=aa''=0$; thus $a'a''=a''a'=0$.

b) We choose ϵ_k such that

$$\text{(i) } Ma' = M\epsilon_k, \quad \text{(ii) } \epsilon_1a^i\epsilon_k = 0, \quad \text{(iii) } a''\epsilon_k = 0, \qquad (i = 0, \cdots, k-2).$$

This is possible for the following reasons. Put $R=\epsilon_1M$ and let u be a nonvanishing element of $R\cap Ma'$. Let $u_1, \cdots, u_p$ be linearly independent elements such that $R\cap Ma''=\mathsf{V}(u_1, \cdots, u_p)$. According to the corollary of Lemma 5 the elements $\epsilon_1, \epsilon_1a, \cdots, \epsilon_1a^{k-2}, u_1, \cdots, u_p, u$ are then linearly independent. The ideal Ma' is simple. According to Lemma 4 there is an ϵ_k in Ma' such that $\epsilon_1a^i\epsilon_k=0$ $(i=0, \cdots, k-2)$, as well as $u_1\epsilon_k= \cdots =u_p\epsilon_k=0$. The left ideal Ma'' is direct sum of the simple left ideals, namely $Mu_1, \cdots, Mu_p$. In particular, we have $a''=s_1u_1+ \cdots +s_pu_p$ for some $s_1, \cdots, s_p$, and thus $a''\epsilon_k=0$. With this choice of ϵ_k we have, therefore, $a'=a'\epsilon_k=a^{k-1}\epsilon_k$, and since $a'=\epsilon_1a^{k-1}$, $a'=\epsilon_1a^{k-1}=a^{k-1}\epsilon_k$.

c) For $\nu=2, \cdots, k-1$ we choose ϵ_ν in such a way that

$$\text{(iv) } M\epsilon_\nu = M\epsilon_1a^{\nu-1}, \qquad \text{(v) } \epsilon_\nu M = a^{k-\nu}\epsilon_kM.$$

This is possible for the following reasons. Let ν take one of the values indicated

above. The left ideal $L=M\epsilon_1 a^{\nu-1}$ as well as the right ideal $R=a^{k-\nu}\epsilon_k M$ are both simple. Let u be a nonvanishing element of $R\cap L$, i.e., $u=p\epsilon_1 a^{\nu-1}=a^{k-\nu}\epsilon_k q$ for some p and q. It follows that $u^2=p\epsilon_1 a^{k-1}\epsilon_k q$. As we have seen before, $\epsilon_1 a^{k-1}\epsilon_k\neq 0$. Since here $p\epsilon_1\neq 0$ as well as $\epsilon_k q\neq 0$, it follows from Lemma 2 that $u^2\neq 0$. Therefore $u^2=\xi u$ with $\xi\neq 0$, and we can put $\epsilon_\nu=(1/\xi)u$.

From this construction we deduce the following statements.

A. *The elements* $\epsilon_1, \cdots, \epsilon_k$ *are pairwise orthogonal.*

Proof. Suppose first that ν takes one of the values $2, \cdots, k-1$. According to (iv) and (v) there are elements x_ν, y_ν such that $\epsilon_\nu=x_\nu\epsilon_1 a^{\nu-1}=a^{k-\nu}\epsilon_k y_\nu$ holds. It follows that $\epsilon_\nu\epsilon_\mu=x_\nu\epsilon_1 a^{k-1+\nu-\mu}\epsilon_k y_\mu$. From $a^k=0$ and (ii) we get thus $\epsilon_\nu\epsilon_\mu=0$ for $\nu\neq\mu$; ν, $\mu=2, \cdots, k-1$. We have $\epsilon_1\epsilon_\nu=\epsilon_1 a^{k-\nu}\epsilon_k y_\nu$; thus $\epsilon_1\epsilon_\nu=0$ for $\nu=2, \cdots, k-1$ (according to (ii)). Again, according to (ii), $\epsilon_1\epsilon_k=0$. For $\nu=2, \cdots, k-1$ we have $\epsilon_\nu\epsilon_k=x_\nu\epsilon_1 a^{\nu-1}\epsilon_k$; therefore, again according to (ii), $\epsilon_\nu\epsilon_k=0$. This proves the orthogonality.

B. $\epsilon_\nu a=a\epsilon_{\nu+1}$ *for* $\nu=1, \cdots, k-1$.

Proof. $\epsilon_\nu a=x_\nu\epsilon_1 a^\nu$ holds for $\nu=2, \cdots, k-1$. According to (iv) we have $x_\nu\epsilon_1 a^\nu\in M\epsilon_{\nu+1}$ for $\nu=1, \cdots, k-2$, thus $\epsilon_\nu a=\epsilon_\nu a\epsilon_{\nu+1}$ for $\nu=2, \cdots, k-2$. According to (iv) we have $M\epsilon_2=M\epsilon_1 a$; thus $\epsilon_1 a=x\epsilon_2=\epsilon_1 a\epsilon_2$. Since $\epsilon_{k-1}a=\epsilon_{k-1}a\epsilon_k$, we have thus $\epsilon_\nu a=\epsilon_\nu a\epsilon_{\nu+1}$ for $\nu=1, \cdots, k-1$. In exactly the same way one shows that $a\epsilon_{\nu+1}=\epsilon_\nu a\epsilon_{\nu+1}$ for $\nu=1, \cdots, k-1$.

C. *If* $e_1=\epsilon_1+\cdots+\epsilon_k$, *then* $e_1 a=ae_1=e_1 ae_1$.

Proof. According to (i) we have $\epsilon_k=xa'$; thus $\epsilon_k a=xa'a=0$ and $e_1 a=\epsilon_1 a+\cdots+\epsilon_{k-1}a$. From $a\epsilon_1=0$ it follows that $ae_1=a\epsilon_2+\cdots+a\epsilon_k$; thus $e_1 a=ae_1=\epsilon_1 a\epsilon_2+\cdots+\epsilon_{k-1}a\epsilon_k$. Since e_1 is idempotent and of rank k we have thus proved the first part of Theorem 1.

D. $e_1 ae_1$ *is maximal nilpotent in* $e_1 Me_1$.

Proof. e_1 is of rank k and we have $e_1 ae_1=\epsilon_1 a\epsilon_2+\cdots+\epsilon_{k-1}a\epsilon_k$. It follows that $(e_1 ae_1)^{k-1}=\epsilon_1 a\epsilon_2\cdots a\epsilon_k$, $(e_1 ae_1)^k=0$. Since $\epsilon_\nu a\epsilon_{\nu+1}\neq 0$ for $\nu=1, \cdots, k-1$ it follows from Lemma 2 that $(e_1 ae_1)^{k-1}\neq 0$.

E. *There is a normal basis of* $e_1 Me_1$ *which is characteristic for* $e_1 ae_1$.

Proof. We construct such a basis, namely $\epsilon_{\nu\mu}(\nu, \mu=1, \cdots, k)$ in the following way: we put $\epsilon_{\nu\nu}=\epsilon_\nu$ $(\nu=1, \cdots, k)$, $\epsilon_{\nu,\nu+1}=\epsilon_\nu a\epsilon_{\nu+1}$ $(\nu=1, \cdots, k-1)$, and complete this basis in the sense of Lemma 3. This basis is then characteristic for $e_1 ae_1$, for in this case part a) of the definition of the characteristic basis has to be applied.

With this the proof of Theorem 1 is completed.

THEOREM 2. *If a is nilpotent then there is a decomposition of e which is characteristic for a.*

Proof. If a is nilpotent of order 1 then any primitive orthogonal decomposition of e is characteristic for a. Suppose now that a is nilpotent of order k_1 with $k_1>1$. We construct e_1 in the sense of Theorem 1. Then $(a-e_1 ae_1)^i=a^i-(e_1 ae_1)^i$; accordingly, $a-e_1 ae_1$ is nilpotent of order k_2 with $k_2\leqq k_1$. If $k_2=1$, and if $\epsilon_1+\cdots+\epsilon_t$ is a primitive orthogonal decomposition of $e-e_1$, then $e=e_1+\epsilon_1+\cdots+\epsilon_t$

is a characteristic decomposition for a, for we have $a=e_1ae_1$ and $e-e_1$ and e_1 are orthogonal. If $k_2>1$ then we consider the total matrix algebra $(e-e_1)M(e-e_1)$. Since $(e-e_1)(a-e_1ae_1)(e-e_1)=a-e_1ae_1$, the element $a-e_1ae_1$ is contained in it. According to Theorem 1, there is an idempotent element e_2 of rank k_2 in this algebra such that $e_2(a-e_1ae_1)=(a-e_1ae_1)e_2$ is nilpotent of order k_2. From $e_2=(e-e_1)e_2(e-e_1)$ it follows that $e_1e_2=e_2e_1=0$, i.e., $e_2(a-e_1ae_1)=e_2a$ and $(a-e_1ae_1)e_2=ae_2$. We have therefore $e_2a=ae_2$ and this element is nilpotent of order k_2. In the same manner we continue with $a-e_1ae_1-e_2ae_2$ and arrive after a finite number of steps at an element $a-e_1ae_1-\cdots-e_sae_s$ which is nilpotent of order 1, i.e., $=0$. If $\epsilon_1+\cdots+\epsilon_t$ is a primitive orthogonal decomposition of $e-e_1-\cdots-e_s$ then the decomposition $e=e_1+\cdots+e_s+\epsilon_1+\cdots+\epsilon_t$ is characteristic for a.

THEOREM 3. *If a is nilpotent in M then there is a basis of M which is characteristic for a.*

Proof. Let $e=e_1+\cdots+e_k$ be a decomposition which is characteristic for a. To every $i(i=1,\cdots,k)$ we construct a basis of e_iMe_i which is characteristic for e_iae_i. The construction of the basis of M characteristic for a is possible according to the corollary of Lemma 3.

COROLLARY. *If a has exactly one eigenvalue then there is a basis of M which is characteristic for a.*

This is trivial, for here $a-\lambda e$ is nilpotent if λ is the eigenvalue of a.

THEOREM 4. *For any element a of M there is a basis of M which is characteristic for it.*

Proof. This follows immediately from the corollary to Lemma 3 if a is represented in the form (1).

References

1. A. A. Albert, Structure of Algebras, Amer. Math. Soc. Coll. Publ., 24(1939) 51.

2. What we are dealing with is nothing else than the ring of residues of all polynomials in x with respect to the minimum polynomial. For the application to matrices see e.g. H. Schwerdtfeger, *Les fonctions de matrices*, ASI, Hermann, Paris, 1938, 649.

A MAXIMIZATION PROBLEM SUGGESTED BY *BAKER* VS. *CARR*

H. E. REINHARDT, University of Montana

1. Legislative Reapportionment. In recent decisions, including the celebrated case of *Baker* vs. *Carr* [369 U.S. 186 (1962)], the United States Supreme Court has invoked a doctrine of "one man, one vote," and ruled that apportionment in state legislatures must be essentially proportional to population. During the anguish and conjecturing about the political effects of this decision that fol-

lowed we were asked if mathematics had anything to add to the discussion. We invoked our own doctrine that in the land of the blind one one-eyed man has as much right to be king as any other and agreed to think about the problem.

Obstacles are immediately apparent in trying to formulate any mathematical problem concerning reapportionment. In the first place, the Supreme Court has not specified what "essential" equality of the values of votes might be. In the second place, reapportionment of an assembly is supposed to be done by the assembly itself, and the quality of a reapportionment scheme for a particular legislator might depend almost entirely on whether or not it has a district he can represent. We have chosen to ignore the second and make some assumptions about the first. According to Paul David and Ralph Eisenberg [1], drafters of model plans of representation have considered the population represented by a legislative member and have suggested that the ratio of largest population per member to smallest population per member should not exceed 1.3. David and Eisenberg, themselves, suggest that 1.5 is politically more feasible.

We can now formulate a problem of possible interest to political bosses and political observers: Among acceptable apportionments, what is the maximum fraction of a legislative assembly which one party can expect to control? We suppose that there are n nonempty voting units (e.g., precincts or counties) in the state. Let p_i be the fraction of the state's population in the ith district. In a perfectly apportioned legislature, this district would be represented by a fraction p_i of the legislators. We suppose that there are only two parties which we call Republican and Democratic out of force of habit. Let r_i, $0 \leqq r_i \leqq 1$, be the fraction of the ith delegation which the Republicans can "expect" to control. (Specification of a reasonable set of r_i's is, of course, difficult; it turns out that it is also the crucial part of the simplified problem we are considering.) We suppose that $d_i = 1 - r_i$ is the fraction the Democrats can expect to control. In that demographic miracle, a perfectly apportioned legislature, the Republicans could expect to control a fraction $\sum_{i=1}^{n} r_i p_i$. (Hereafter we will suppress these limits of summation.) The apportionment of any legislature can be represented by a set of positive numbers $\{w_i\}$ such that the fraction of the assembly representing the ith voting unit is $w_i p_i / \sum w_i p_i$. (The w_i for instance might be ratio of number of legislators to population.) If the set $\{w_i\}$ represents an apportionment the set $\{k w_i\}$ also represents that apportionment for any positive k. The fraction of the assembly controlled by the Republicans is $\sum w_i r_i p_i / \sum w_i p_i$. In terms of this notation an acceptable apportionment, according to the David-Eisenberg criterion, is $\max w_i / \min w_i \leqq 1.5$. The problem of interest to bosses and observers can now be posed as follows: Among all apportionments with $\max w_i / \min w_i \leqq 1.5$ what are maximum and minimum values of $\sum w_i r_i p_i / \sum w_i p_i$? We note that we may take $\min w_i = 1$ so the restraint can be taken as $1 \leqq w_i \leqq 1.5$.

2. A solution via the Neyman-Pearson Lemma. The solution of the problem is given by application of the following

THEOREM. *Let $f_1, f_2, \cdots, f_n, g_1, g_2, \cdots, g_n, A$, and B be $2n+2$ real numbers with the f_i nonnegative, the g_i positive, and A and B satisfying $0<A<B$. The function of n variables*

$$\psi(w_1, w_2, \cdots, w_n) = \sum w_i f_i / \sum w_i g_i$$

has a maximum ψ^* on the hypercube H in Euclidean n-space defined by $A \leqq w_i \leqq B$, $1 \leqq i \leqq n$; $\psi(w_1, w_2, \cdots, w_n) = \psi^*$ if and only if

$$\begin{aligned} w_i &= A \text{ if } f_i/g_i < \psi^* \\ w_i &= B \text{ if } f_i/g_i > \psi^*. \end{aligned} \tag{1}$$

Proof. Since $\psi(x)$ is a continuous function defined on a compact set it attains its least upper bound. Call it ψ^*. In the search for the maximum we use the following slight modification of the Neyman-Pearson Lemma of statistical hypothesis testing:

LEMMA. *Let $f_1, \cdots, f_n, g_1, \cdots, g_n, A$, and B be $2n+2$ nonnegative numbers with $A<B$. Let H be the hypercube in Euclidean n-space defined by $A \leqq w_i \leqq B$. For any c satisfying $A \sum g_i \leqq c \leqq B \sum g_i$, the quantity $\sum f_i w_i$ is maximized over H subject to the restraint $\sum g_i w_i = c$ by a point $(w_1, \cdots, w_n)$ with*

$$\begin{aligned} w_i &= B \text{ if } f_i/g_i > k \\ w_i &= A \text{ if } f_i/g_i < k, \end{aligned} \tag{2}$$

where k and the values of w_i for those i with $f_i/g_i = k$ are chosen to satisfy the restraint. All maximizing solutions satisfy (2), *for the same k.*

In many books (e.g. [**2**], pp. 294–6) the Neyman-Pearson Lemma is a result concerning the choice of a set in a probability space. But the specification of a set is equivalent to the specification of its indicator (or, synonymously, characteristic) function—a function with range $\{0, 1\}$. One can proceed by easy steps to replace functions with the two-point range to functions whose range is the entire closed interval $[0, 1]$. (The details appear, for instance, on pages 65 and 66 of [**3**].) It is one more easy step to a function whose range is an arbitrary interval $[A, B]$. Trivial modifications of the proof of [**3**] suffice.

To prove the theorem we note that at a point $(w_1, w_2, \cdots, w_n)$ for which $\psi(w_1, w_2, \cdots, w_n) = \psi^*$ one must have $\sum g_i w_i = c$ for some c. If one knew this value of c he could replace the problem of maximizing ψ by the problem of maximizing $\sum f_i w_i / c$ over S subject to the restraint $\sum g_i w_i = c$. The solution to this problem is given by the Neyman-Pearson Lemma. But since we know that $\sum g_i w_i = c$ for some c we know that a maximizing point is given by (2) for some k.

Suppose that $(w_1, \cdots, w_n)$ satisfies (2) but not (1) and that $\psi(w_1, \cdots, w_n) = \psi^*$. Then for some i, say i_0, $f_{i_0}/g_{i_0} < \psi^*$ and $w_{i_0} > A$ or $f_{i_0}/g_{i_0} > \psi^*$ and $w_{i_0} < B$. But, from the fact that (for b, y, and $t-1$ positive)

$$\frac{a+x}{b+y} - \frac{a+tx}{b+ty}$$

has the same sign as

$$\frac{a+x}{b+y}-\frac{x}{y}$$

we see that replacing w_{i_0} by A in the first case or w_{i_0} by B in the second will give a point in the set H with $\psi(w_1, \cdots, w_n) > \psi^*$. This is a contradiction and completes the "only if" part of the proof. The "if" part follows from the existence of the maximum and the fact that if

$$\frac{a+x}{b+y}=\frac{x}{y},$$

then for all $t>0$,

$$\frac{a+tx}{b+ty}=\frac{x}{y}.$$

The maximum Republican representation for the problem of Section 1 is given by the theorem with $A=1$, $B=1.5$, $g_i=p_i, f_i=r_ip_i$. For maximum Democratic representation, we would, of course, take $f_i=d_ip_i$, and obtain from this the minimum Republican representation. We alluded earlier to the crucial nature of the choice of the r_i's. For one reasonable specification of the r_i's the maximum and minimum Republican representations for the Montana legislature were .65 and .50 while for another equally reasonable specification they were .40 and .23.

3. Another application of the theorem. While it is unlikely that the solution sheds much light on the serious matter of reapportionment, the theorem of Section 2 has another application. In their paper [4], Ostrowski and Schneider prove and use the theorem for the case $g_i=1$, $i=1, 2, \cdots, n$ to obtain bounds on the maximal characteristic root of a nonnegative irreducible matrix. Their proof uses a mean-value theorem where we have used the Neyman-Pearson Lemma.

The problem considered by Ostrowski and Schneider has a simple physical interpretation. To each of n fixed points on a line with one-dimensional co-ordinates $f_1, f_2, \cdots, f_n$ assign a mass w_i, $A \leqq w_i \leqq B$, $i=1, 2, \cdots, n$ in such a way as to maximize the co-ordinate of the center of gravity. A small set of physicists, confronted with the problem expressed in these terms, with $A=\epsilon$, $B=1$, conjectured that the solution would be $w_i=1$ if and only if $f_i=\max_j f_j$ and $w_i=\epsilon$ otherwise. This, one of them supposed, because of an instinctive faith in the smallness of epsilons. (An equally small set of political observers confronted with the problem of Section 1 conjectured that the maximum would occur if $w_i=1.5$ if and only if $r_i>.5$ and $w_i=1$ otherwise. What sort of instinctive faith that represents is not clear.)

Acknowledgement: Professor H. Schneider suggested that the Neyman-Pearson Lemma might prove useful in the problem mentioned in Section 3. My colleague, Professor Ellis Waldron of the U.M. political science department raised the problem of Section 2. The author is grateful to both.

References

1. Paul T. David and Ralph Eisenberg, Devaluation of the urban and suburban vote, Bureau of Public Administration of Virginia, Charlottesville, 1961.

2. Harold Freeman, Introduction to Statistical Inference, Addison-Wesley, Reading, Mass. 1963.

3. E. L. Lehmann, Testing Statistical Hypotheses, Wiley, New York, 1959.

4. A. Ostrowski and H. Schneider, Bounds for the maximal characteristic root of a non-negative irreducible matrix, Duke Math. J., 27 (1960) 547–553.

REGULARITY OF CERTAIN EXPLICIT SOLUTIONS TO LAPLACE'S EQUATION AT ARTIFICIAL INTERFACES

R. B. KELMAN, University of Maryland and Howard University
(Now at Colorado State University)

1. Introduction. Weak solutions to elliptic partial differential equations and their relation to classical solutions have become standard topics in research [5]. It is worthwhile therefore to have available as examples simple problems of physical and numerical interest in which these concepts arise in a natural way. In this note one such problem is discussed in detail. I have found that the analysis given here and the elegantly easy proof of Weyl's lemma given in [4] are useful supplements to a standard course (e.g. [2]) dealing with explicit methods of solving partial differential equations. This material helps bridge the gap between the classical method of separation of variables and more modern ideas. Because of the simplicity of the explicit formulas involved all proofs can be given rigorously and need not be deferred until a later study of Green's functions and distributions.

The theorem given in Section 2 establishes the regularity of an explicit axisymmetric solution of Laplace's equation. Although the analysis is routine to the best of the writer's knowledge no proof of this regularity has been given heretofore even though the formulas themselves were known at least by 1904 [8]. We confine ourselves to this simple example to make the computations as easy as possible. The proof itself is easily modified, in an almost verbatim way, to deal with other similar problems.

For example, it applies to the problem which arises from the device described in Section 3 for finding an alternative form for the solution of a boundary value problem for the two dimensional Laplace equation. This alternative form provides rapid convergence for values of the independent variables for which the standard series solution converges slowly. In [3, p. 270] another alternative form of a more general nature is given for this problem.

The word "artificial" in the title refers to the fact that, in the problems discussed below, the interface arises as an artifact of the method of solution and not as a result of two dissimilar physical bodies being brought into contact.

2. A steady temperature problem. Consider the steady temperature in an infinite uniform cylinder of radius 1 whose axis coincides with the x-axis. For $x<0$ the surface temperature is 1, and for $x>0$ it is 0. Thus we seek an axisymmetric solution, say $u(x, r)$, of the equation

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} = 0, \qquad -\infty < x < \infty,\ 0 < r < 1 \tag{1}$$

subject to the boundary conditions

$$\begin{aligned} u(x, r) &= 1, && x < 0,\ r = 1 \\ &= 0, && x > 0,\ r = 1. \end{aligned} \tag{2}$$

Following [1, p. 209; 8] a solution of the form

$$\begin{aligned} u(x, r) &= 1 + \sum_{n=1}^{\infty} j_n \mathcal{J}_0(\gamma_n r) e^{\gamma_n x}, && x \leqq 0,\ 0 < r < 1, \\ &= \sum_{n=1}^{\infty} p_n \mathcal{J}_0(\gamma_n r) e^{-\gamma_n x}, && x \geqq 0,\ 0 < r < 1, \end{aligned} \tag{3}$$

is sought where γ_n is the nth positive zero of $J_0(r)$, a Bessel function of the first kind, and $\mathcal{J}_0(\gamma_n r) = \sqrt{2} J_0(\gamma_n r)/J_1(\gamma_n)$ is a normalized Bessel function. The values of j_n and p_n are formally determined by assuming that u and $\partial u/\partial x$ are continuous at $x=0$ and matching coefficients in the Fourier-Bessel series given in (3). Remembering that $1 = \sum \sqrt{2}\gamma_n^{-1} \mathcal{J}_0(\gamma_n r)$ we find that

$$j_n + \frac{\sqrt{2}}{\gamma_n} = p_n, \qquad j_n = -p_n.$$

Thus (3) yields

$$\begin{aligned} u(x, r) &= 1 - \sum_{n=1}^{\infty} \frac{e^{\gamma_n x}}{\sqrt{2}\gamma_n} \mathcal{J}_0(\gamma_n r), && x \leqq 0,\ 0 < r < 1, \\ u(x, r) &= \sum_{n=1}^{\infty} \frac{e^{-\gamma_n x}}{\sqrt{2}\gamma_n} \mathcal{J}_0(\gamma_n r), && x \geqq 0,\ 0 < r < 1. \end{aligned} \tag{4}$$

Set $R_1^0 = \{x<0;\ 0<r<1\}$, $R_2^0 = \{x=0;\ 0<r<1\}$, $R_3^0 = \{x>0;\ 0<r<1\}$ and $R^0 = R_1^0 \cup R_2^0 \cup R_3^0$. Clearly, u is continuous in R^0, harmonic in $R_1^0 \cup R_3^0$, and satisfies the boundary condition (2). It remains to be shown that u is harmonic in the neighborhood of R_2^0 and hence throughout R^0.

REMARK. If u_x and u_r were continuous on R_2^0 the problem would be solved at

once by the theorem on analytic continuation [6, p. 261], but from the explicit form (4) no information is available concerning the behavior of u_x or u_r on R_2^0. Intuitively speaking, we have "weakly matched" the left and right branches of u_x along R_2^0, and we now wish to show that this implies "strong (i.e. pointwise) matching" along R_2^0.

Let us recall that a weak solution of (1) is a function, say v, locally Lebesgue integrable with respect to the weight function r and such that

$$\iint_D v\Delta f r dr dx = 0$$

whenever D is a compact subset of R^0 and f is an infinitely differentiable function that vanishes exterior to an open subset of D. Further, if v is a weak solution then by Weyl's Lemma v is equal almost everywhere to a harmonic function [4; 5]. Since u is continuous, it is harmonic if it is a weak solution. Therefore let us prove the

LEMMA. *The function $u(x, r)$ given by* (4) *is a weak solution of* (1).

Proof. Since u is harmonic in $R_1^0 \cup R_3^0$ it follows by Green's second identity [6, p. 215] that u is a weak solution of (1) if

$$\iint_D u\Delta f r dr dx = 0$$

whenever D is a rectangle of the form

$$D = \{-a \leqq x \leqq a; b \leqq r \leqq c\},$$

where a, b, and c are positive constants such that $c<b<1$ and f is an infinitely differentiable function vanishing exterior to an open subset of D. Let δ be a small positive number. Let A_1 (see Fig. 1) be the area of D to the left of $x=-\delta$, A_2 the area between $x=-\delta$ and $x=\delta$, and A_3 the area to the right of $x=\delta$. Denote

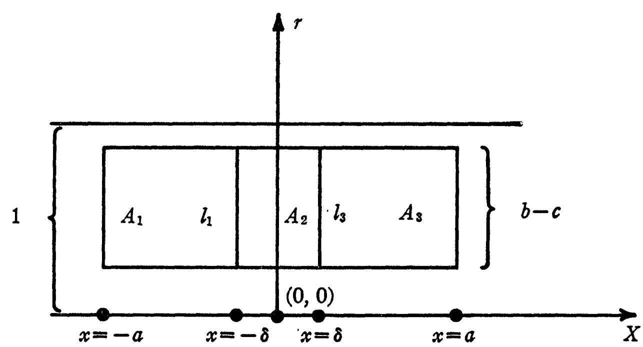


FIG. 1. An illustration in the x-r plane of the area D used in proving that u is a weak solution.

by l_1 and l_3 respectively the segments of $x=-\delta$ and $x=\delta$ intercepted by D. Clearly

$$\lim_{\delta=0}\int\int_{A_2} u\Delta f r dr dx = 0.$$

Using Green's second identity and remembering that u is harmonic in $R_1^0\cup R_3^0$ we have

$$\int\int_{A_1\cup A_3} u\Delta f r dr dx = \int_{l_3} f\frac{\partial u}{\partial x} r dr - \int_{l_1} f\frac{\partial u}{\partial x} r dr + \int_{l_1} u\frac{\partial f}{\partial x} r dr - \int_{l_3} u\frac{\partial f}{\partial x} r dr. \tag{5}$$

Now

$$\begin{aligned}\int_{l_3} f\frac{\partial u}{\partial x} r dr - \int_{l_1} f\frac{\partial u}{\partial x} r dr &= \int_0^1 f(\delta, r)\left(\frac{\partial u}{\partial x}\bigg|_{x=\delta} - \frac{\partial u}{\partial x}\bigg|_{x=-\delta}\right) r dr \\ &\quad + \int_0^1 (f(\delta,r)-f(-\delta,r))\frac{\partial u}{\partial x}\bigg|_{x=-\delta} r dr.\end{aligned} \tag{6}$$

Since $u_x|_{x=\delta}=u_x|_{x=-\delta}$, the first integral on the right hand side of (6) vanishes. Define $F(x, r)$ by $F(x, r)=f(x, r)-f(-x, r)$. Now F is uniformly continuous in D, and in particular

$$\lim_{x=0} F(x, r) = 0, \qquad 0 \leqq r \leqq 1 \tag{7}$$

uniformly in r. Let us expand F in a Fourier-Bessel series

$$F(x, r) = \sum_{n=1}^{\infty} F_n(x)\mathcal{J}_0(\gamma_n r),$$

where $F_n(x)=\int_0^1 F(x, r)\mathcal{J}_0(\gamma_n r) r dr$. Employing integration by parts twice we see that

$$F_n(x) = -\frac{1}{\gamma_n^2}\int_0^1 \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial F}{\partial r}\right)\mathcal{J}_0(\gamma_n r) r dr.$$

Now $r^{-1}(rF_r)_r$ is square integrable with respect to the weight r on the interval $0<r<1$ for any choice of x. By the Riesz-Fischer Theorem [7] its Fourier-Bessel coefficient is $O(n^{-1/2})$ as $n\to\infty$. Consequently, $F_n(x)=O(n^{-5/2})$ since $\gamma_n^{-1}=O(n^{-1})$. Using Parseval's identity [7] we see that

$$\left|\int_0^1 \frac{\partial u}{\partial x}\bigg|_{x=-\delta} F(\delta, r) r dr\right| = \left|\frac{1}{\sqrt{2}}\sum_{n=1}^{\infty} e^{-\gamma_n\delta}F_n(\delta)\right| \leqq \frac{1}{\sqrt{2}}\sum_{n=1}^{\infty}\left|F_n(\delta)\right|. \tag{8}$$

From (7) $\sum_n F_n^2(x)\to 0$ as $x\to 0$. Hence $F_n(x)\to 0$ as $x\to 0$ uniformly in n. Therefore the last series in (8) tends to 0 as $\delta\to 0$. In a similar way one can show that the sum of the two remaining integrals on the right hand side of (5) tends to 0 as $\delta\to 0$. This proves the lemma. In summary we have shown

THEOREM. *The function u given by* (4) *is a twice continuously differentiable solution of* (1) *satisfying the boundary conditions* (2).

3. Discussion. Carslaw and Jaeger [1] and Wilson [8] give other examples of this type, e.g. the steady temperature in a cylinder moving in the direction of its axis and satisfying (2). Moreover, the type of solution discussed above can be useful when alternate solutions are easily available. We give an example: to find the steady temperature in a uniform unit square, i.e.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad 0 < x < 1, 0 < y < 1,$$

with the boundary conditions

$$\begin{aligned} u(x, y) &= 1, & x &= 0, 0 < y < 1 \\ u(x, y) &= 1, & 0 &< x < l, y = 0, 1, \\ u(x, y) &= 0, & l &< x < 1, y = 0, 1, \\ u(x, y) &= 0, & x &= 1, 0 < y < 1, \end{aligned}$$

where l is a constant such that $0 < l < 1$.

From [1, Ch. V] the solution is

$$\begin{aligned} u(x, y) = {} & \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(1 - \cos n\pi l) \sin n\pi x(\sinh (1 - y)n\pi + \sinh n\pi y)}{n \sinh n\pi} \\ & + \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sin (2n + 1)\pi y(\sinh (2n + 1)\pi(1 - x))}{(2n + 1) \sinh (2n + 1)\pi}, \\ & \qquad 0 < x < 1, 0 < y < 1. \end{aligned} \tag{9}$$

On the other hand matching across an artificial interface at $\{x = l;\ 0 < y < 1\}$ an alternate form for the above solution is

$$\begin{aligned} u(x, y) = {} & 1 \\ & - \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(\tanh (2n + 1)\pi l)(\sin (2n + 1)\pi y) \sinh (2n + 1)\pi x}{(2n + 1)(\tanh (2n + 1)\pi l + \tanh (2n + 1)\pi(1 - l)) \sinh (2n + 1)\pi l}, \\ & \qquad 0 < x \leqq l; 0 < y < 1, \\ u(x, y) = {} & \frac{4}{\pi} \\ & \cdot \sum_{n=0}^{\infty} \frac{\tanh (2n + 1)\pi(1 - l)(\sin (2n + 1)\pi y) \sinh (2n + 1)\pi(1 - x)}{(2n + 1)(\tanh (2n + 1)\pi l + \tanh (2n + 1)\pi(1 - l)) \sinh (2n + 1)\pi(1 - l)}, \\ & \qquad l < x < 1, 0 < y < 1. \end{aligned} \tag{10}$$

For numerical purposes the series in (9) converge rapidly provided x is not too

close to 0, or y is not too close to 0 or 1, because of the factors sinh $n\pi y$/sinh $n\pi$ etc. When x or y are near these values the series in (9) converge slowly and present difficulties especially in trying to write a simple program for a computor. On the other hand the series in (10) converge rapidly for these values of x and y but they do not converge rapidly near $x=l$. Thus the two representations of the solution can be used for effective and easy numerical calculations over the whole domain of the independent variables, whereas difficulties arise in using only one or the other of the formulas.

Introducing an artificial interface can be applied in a straightforward manner to other similar boundary value problems to obtain alternative forms of the solution.

The writer thanks Mr. Y. H. Lee for having read this manuscript critically.

This work was supported in part by a contract of the Office of Naval Research, U. S. Navy, Contract No. 595 (17) with the University of Maryland.

References

1. H. S. Carslaw and J. C. Jaeger, Conduction of Heat in Solids, Oxford, 1959.
2. R. V. Churchill, Fourier Series and Boundary Value Problems, 2nd ed. McGraw-Hill, New York, 1963.
3. B. Friedman, Principles and Techniques of Applied Mathematics, Wiley, New York, 1956.
4. L. Gårding, On a Lemma by H. Weyl, Kungl. Fysiogr. Sällsk. i Lund Förh., 20 (1951) 1–4.
5. G. Hellwig, Partielle Differentialgleichungen, Stuttgart, 1960.
6. O. D. Kellogg, Foundations of Potential Theory, Dover reprint, New York, 1953.
7. F. Riesz, and B. Sz.-Nagy, Leçons d'Analyse Fonctionnelle, Budapest, 1952.
8. H. A. Wilson, On Convection of Heat, Proc. Cambridge Philos. Soc., 12 (1904) 406–423.

MAPPINGS OF SEMIGROUPS ASSOCIATED WITH ORDERED PAIRS

MARGARET L. VITANZA, SUNY at Buffalo

I. Introduction. In 1961, S. P. Franklin and J. W. Lindsay, in [3], defined a straddle on a semigroup S to be an ordered pair (a, b) of elements of S with the property that $axb=x$ for all x in S. This notion of straddle leads in a natural way to the examination of mappings defined on S which carry an arbitrary element x of S into the element axb, where a and b are fixed elements of S. Specifically, for an ordered pair (a, b), we define a mapping ${}_aT_b$ from S into S by ${}_aT_b(x)=axb$ and find necessary and sufficient conditions that ${}_aT_b$ be each of the following: epimorphism, automorphism, monomorphism, anti-epimorphism and anti-automorphism.

Let us recall that a mapping ϕ from S into S is an endomorphism if $\phi(x)\phi(y)=\phi(xy)$ for all x, y in S. If ϕ is onto, ϕ is said to be an epimorphism; and if, in addition, ϕ is one-one, then ϕ is called an automorphism. An endomorphism which is one-one is referred to as a monomorphism. We say that ϕ is an anti-

endomorphism if $\phi(x)\phi(y)=\phi(yx)$ for all x, y in S. The terms anti-epimorphism and anti-automorphism are then defined analogously. Quite often, when no confusion can result, we denote ${}_aT_b$ more simply by T. Our main result is contained in Theorem 1, which shows that in order that ${}_aT_b$ be an epimorphism *or* automorphism for some ordered pair (a, b), it is necessary and sufficient that S have an identity e, and $ab=ba=e$. This result is then used to obtain the theorem of Franklin and Lindsay, namely that (a, b) is a straddle on S if and only if S has an identity e, $ab=ba=e$, and a and b are in the center of S.

II. Theorems and proofs.

THEOREM 1. *The following statements are equivalent*:

(1) ${}_aT_b$ *is an epimorphism*,

(2) S *has an identity* $e=ab=ba$,

(3) ${}_aT_b$ *is an automorphism.*

Proof. (1) implies (2). Consider x in S. Since T is onto, there exists a y such that $x=T(y)$. Then, we have $T(a)x=T(a)T(y)=T(ay)=aayb=aT(y)=ax$; that is,

$$aabx = ax. \tag{1.1}$$

Also, $xT(b)=T(y)T(b)=T(yb)=aybb=T(y)b=xb$; that is,

$$xabb = xb. \tag{1.2}$$

Then, $[T(ab)]x=T(ab)T(y)=T(aby)=aabyb=(aaby)b=ayb=x$ by (1.1). Similarly, $xT(ab)=T(y)T(ab)=T(yab)=ayabb=a(yabb)=ayb=x$ by (1.2). Hence $xT(ab)=T(ab)x=x$ and $T(ab)=e$ is the identity for S. Furthermore, $T(ab)=aabb=ab$ by (1.2) and thus $ab=e$. We also have $T(ba)=T(b)T(a)$; that is, $abab=abbaab$. But since $ab=e$, $e=abab=(ab)ba(ab)=ba$.

(2) implies (3): $T(x)T(y)=axbayb=ax(ba)yb=axyb=T(xy)$. Thus T is an endomorphism.

Now let x in S be given. We wish to find a y in S such that $T(y)=x$. Letting $y=bxa$, we see that $ayb=a(bxa)b=(ab)x(ab)=x$.

If $T(x)=T(y)$; that is, $axb=ayb$, then $b(axb)a=b(ayb)a$ or $(ba)x(ba)=(ba)y(ba)$. But since $ba=e$ by hypothesis, $x=y$.

It is evident that (3) implies (1).

COROLLARY 2. (a, b) *is a straddle on* S *if and only if* S *has an identity* e, $ab=e$, *and* a *and* b *are in the center of* S.

Proof. Suppose that (a, b) is a straddle on S. Then the mapping ${}_aT_b$ is the identity mapping which is an automorphism. Hence, by Theorem 1, $ab=ba$ is the identity for S. Furthermore, $xa=T(xa)=axab=axe=ax$, and $bx=T(bx)=abxb=exb=xb$.

Conversely, if $ab=e$ and a and b are in the center of S, $T(x)=axb=abx=x$ and hence (a, b) is a straddle.

Next we proceed to the investigation of mappings of the form ${}_aT_b$ which are monomorphisms. We consider this problem only for semigroups with identity.

THEOREM 3. *Let S be a semigroup with identity. Then ${}_aT_b$ is a monomorphism for some (a, b) if and only if $ba=e$.*

Proof. Suppose T is a monomorphism. Then $T(xy)=T(x)T(y)$; that is, $axyb=axbayb=a(xbay)b$. Since T is one-one, we have $xy=xbay$. Letting $x=y=e$ in the preceding equation, we have $e=ba$.

On the other hand, if $ba=e$, it follows that T is a monomorphism, using techniques similar to those used in the proof of Theorem 1.

We now turn to the investigation of conditions under which T is an anti-epimorphism and anti-automorphism. We are able to obtain the following results:

THEOREM 4. *The following statements are equivalent:*
(1) *${}_aT_b$ is an anti-epimorphism,*
(2) *S is commutative, has an identity e and $ab=e$,*
(3) *${}_aT_b$ is an anti-automorphism.*

Proof. (1) implies (2): Consider x in S. There exists a y such that $x=T(y)$, and $xT(a)=T(y)T(a)=T(ay)=aayb=aT(y)=ax$; that is,

$$xaab = ax. \tag{4.1}$$

We also have $T(b)x=T(b)T(y)=T(yb)=aybb=T(y)b=xb$ which implies

$$abbx = xb. \tag{4.2}$$

$[T(ba)]x=T(ba)T(y)=T(yba)=aybab=[T(y)]ab$ and therefore

$$[T(ba)]x = xab. \tag{4.3}$$

Using (4.2), we get $x(ba)=(xb)a=(abbx)a=(abb)xa=x(ab)$. Thus

$$x(ba) = x(ab). \tag{4.4}$$

In a similar manner, we use (4.1) to obtain $(ba)x=b(ax)=b(xaab)=bx(aab)$ $=a(bx)$. Hence,

$$(ba)x = (ab)x. \tag{4.5}$$

Now $(ab)x=(baab)x$ by (4.1). But $(baab)x=ba(abx)=ab(abx)=(abab)x$ $=[T(ba)]x=x(ab)$ by (4.3) and (4.4). Therefore, we have

$$(ab)x = x(ab). \tag{4.6}$$

Similarly,

$$(ba)x = x(ba) \tag{4.7}$$

since $(ba)x=(ab)x=x(ab)=x(ba)$ by (4.4) and (4.5).

We note that $axb = abx$ since $(ab)x = x(ab) = x(baab) = xb(aab) = axb$ by (4.1) and (4.5). Therefore, $(ab)x = abT(y) = ab(ayb) = ab(aby) = (abab)y = T(ba)y = y(ab) = (ab)y = ayb = x$ by (4.3) and (4.5).

Hence, $(ab)x = x(ab) = x(ba) = (ba)x$ and $ab = ba = e$ is the identity for S.

Since T is an anti-epimorphism, we have $axyb = aybaxb$ for all x, y in S. Then $b(axyb)a = b(aybaxb)a$; that is, $(ab)xy(ba) = (ba)y(ba)x(ba)$ and hence $xy = yx$.

(2) implies (3): This follows immediately from Theorem 1, since if S is commutative, every automorphism is an anti-automorphism.

It is evident that (3) implies (1).

Our remaining problem is the determination of necessary and sufficient conditions that ${}_aT_b$ be an endomorphism. We are able to obtain the following results:

THEOREM 5. *Let S be a commutative semigroup with identity. Then ${}_aT_b$ is an endomorphism if and only if ab is the identity of $T[S]$, where $T[S] = \{T(x): x \in S\}$.*

Proof. Necessity: Suppose that T is an endomorphism and x is any element of $T[S]$. Then there exists a y in S such that $T(y) = x$.

$$[T(e)]x = T(e)T(y) = T(ey) = T(y) = x.$$

Similarly, $xT(e) = T(y)T(e) = T(ye) = T(y) = x$. But $T(e) = aeb = ab$ and hence $ab \in T[S]$ and is the identity of $T[S]$.

Sufficiency: Suppose that ab is the identity of $T[S]$. Then $T(x)T(y) = axbayb = (axyb)ab = axybe = axyb = T(xy)$.

THEOREM 6. *Let S be a cancellation semigroup. Then ${}_aT_b$ is an endomorphism if and only if S has an identity e and $ba = e$.*

Proof. Necessity: Suppose that T is an endomorphism. Then $axyb = axbayb$ for all x, y in S. Since S is cancellative, we have $xyb = xbayb$ and hence $xy = xbay$. Then $y = bay$ and $x = xba$. Therefore, for any x in S, we have $x = bax = xba$ and ba is the identity for S.

Sufficiency: Now suppose that $ba = e$. Then $T(x)T(y) = axbayb = axeyb = axyb = T(xy)$.

III. Some examples and remarks. A reasonable question that arises at this point is whether there exist monomorphisms of the form ${}_aT_b$ which are not automorphisms. Then, by Theorem 3, we are looking for elements a and b, where b is a left inverse of a but not the inverse of a. We note that if a semigroup S contains such elements a and b, then the semigroup contains infinitely many elements. (This statement follows from the fact that if S is finite and $ba = e$, the mapping ${}_aT_b$ is a monomorphism. But since any one-one mapping of a finite set into itself must necessarily be onto, ${}_aT_b$ is an automorphism, and hence $ab = ba = e$ by Theorem 1.) Before we present an example of such a mapping, let us consider the semigroup $\mathfrak{R}(X)$ of all relations on a set X with binary operation defined by $f \circ g = \{(x, y): \exists z \ni (x, z) \in g \text{ and } (z, y) \in f\}$ for f, g in $\mathfrak{R}(X)$. The identity for

$\mathcal{R}(X)$ is the relation $I=\{(x, x): x \in X\}$. Now for any element f in $\mathcal{R}(X)$, we define $f_*=\{(x, y): (y, x) \in f\}$. One can then verify the following statements:

1) Let f be an element of $\mathcal{R}(X)$. Then $f \circ f_* = I$ if and only if

(a) Range of f is X; (b) f is a function.

2) Let f be an element of $\mathcal{R}(X)$. Then $f \circ f_* = f_* \circ f = I$ if and only if

(a) Range of f and f_* is X; (b) f is a one-one function.

Thus, in order to exhibit an example of a mapping T which is a monomorphism but not an automorphism, select ${}_{f_*}T_f$ where f is a function on an infinite set X, the range of f is X, and f is *not* a one-one function.

Next, let S_1 be the closed interval $[0, 1]$ and for x and y in S_1, define $x \circ y = \min\{x, y\}$. Then S_1 is a semigroup with respect to $(\circ)$ and the identity for S_1 is the real number 1. We note that there exists no epimorphism of the form ${}_aT_b$ which is not the identity map and further, that if ${}_aT_b$ is an epimorphism, then $a=b=1$.

Consider a set S_2 with binary operation defined by $x \circ y = y$ for all x, y in S_2. Then S_2 is a semigroup with respect to $(\circ)$. We observe that if S_2 has more than one element, the mapping T is not an epimorphism for any (a, b), for given any $x \neq b$, there exists no y such that $ayb=x$, since $ayb=b$ for all y.

And now a few words about the general problem: As Theorem 6 shows, the problem of determining necessary and sufficient conditions that ${}_aT_b$ be an endomorphism is handled rather easily when the semigroup in question satisfies both cancellation laws. (For a further study of endomorphisms of such semigroups, see [**4**].) However, we have not been able to solve the problem for semigroups in general. Yamada in [**5**] defines an element u of a semigroup S to be a "middle unit of S" if $xuy=xy$ for all x, y in S. Now ${}_aT_b(x)$ is an endomorphism if ba is a middle unit of S. However, the following example shows that one cannot expect the converse to hold; that is, ${}_aT_b$ may be an endomorphism even though ba is not a middle unit of S: Let H be a nonempty set and H^* the set of all subsets of H. For A, B in H^*, define $A \circ B = A \cup B$. Then H^* together with the binary operation of set union is a semigroup. We assert that T is an endomorphism for any pair of elements A, B of H^*, but BA is a middle unit only if $A=B=\varnothing$.

The author would like to express her sincere appreciation to Professor K. D. Magill for his helpful suggestions during the preparation of this paper, I also thank the referee for his valuable comments and for pointing out some of the references mentioned above.

References

1. R. H. Bruck, A Survey of Binary Systems, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1958.

2. A. H. Clifford and G. B. Preston, Algebraic Theory of Semigroups, vol. 1, Amer. Math. Soc., 1961.

3. S. P. Franklin and J. W. Lindsay, Straddles on semigroups, Math. Mag., 34 (1961) 269–270.

4. R. Croisot, Automorphismes intérieurs d'un semi-groupe, Bull. Soc. Math. France, 82 (1951) 161–194.

5. M. Yamada, A note on middle unitary semigroups, Kōdai Math. Sem. Rep., 7 (1955) 49–52.

INEQUALITIES FOR SIMPLEXES

HYMAN GABAI, University of Illinois, UICSM

1. Introduction. In [1] Carlitz derived a number of interesting inequalities related to a point in a triangle. In this paper we consider a point and an n-dimensional simplex. We derive some inequalities with respect to a parameter t which varies according to the position of the point. If the point is in the simplex we may eliminate the parameter in some of the inequalities.

We shall utilize some of the notation and methods which are introduced in a [three-dimensional] geometry course currently being developed by UICSM under the direction of Max Beberman and Herbert Vaughan. The algebraic methods used in that course extend in a natural way to n-dimensions. If X, Y, and Z are points, the vector which maps X onto Y is denoted by $Y-X$, and the image of Z under that mapping is denoted by $Z+(Y-X)$. See [3].

Let A_i $[0\leqq i\leqq n]$ be the vertices of an n-dimensional simplex S_n in n-dimensional space E_n with $n\geqq 2$. $E(j)$ will denote the $n-1$ dimensional subspace containing the points $A_i[0\leqq i\leqq n,\ i\neq j]$. Let $\mathbf{a}_i=A_i-A_0[1\leqq i\leqq n]$, and

$$P = A_0 + \sum_{i=1}^{n} \mathbf{a}_i a_i$$

such that P is not a vertex of S_n and such that the line A_iP intersects $E(i)$ at the point $U_i[0\leqq i\leqq n]$.

Let u_i be the measure of the segment PU_i, and x_i the measure of the segment A_iP. We choose u_i to be nonnegative, and x_i to be positive if A_iP has the same sense as PU_i or if $u_i=0$, and negative otherwise.

We shall derive the relations between the u_i's and x_i's which are stated in the following theorems.

LEMMA.

$$(1)\qquad \frac{u_0}{x_0} = \frac{1-\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} a_i}, \qquad\qquad (2)\qquad \frac{u_i}{x_i} = \frac{a_i}{1-a_i}, \quad [1 \leqq i \leqq n].$$

COROLLARY 1.

$$(3)\qquad \sum_{i=0}^{n} \frac{u_i}{u_i+x_i} = 1, \qquad\qquad (4)\qquad \sum_{i=1}^{n} \frac{u_i}{u_i+x_i} = \sum_{i=1}^{n} a_i.$$

THEOREM. *If* $a_i\geqq 0[1\leqq i\leqq n]$ *and* $0<\sum_{i=1}^{n} a_i\leqq t$ *then*

$$(5)\qquad \prod_{i=1}^{n} (u_i + x_i) \geqq \left(\frac{n}{t}\right)^n \prod_{i=1}^{n} u_i,$$

(6) $$\sum_{i=1}^{n} \frac{x_i}{u_i} \geqq n\left(\frac{n}{t} - 1\right) \qquad [u_i \neq 0]$$

(7) $$\sum_{i=1}^{n} x_i > \left(\frac{1}{t} - 1\right) \sum_{i=1}^{n} u_i.$$

(8) *If* $t \leqq 1$ *then* $$\prod_{i=1}^{n} x_i \geqq \left(\frac{n}{t} - 1\right)^n \prod_{i=1}^{n} u_i.$$

Equality holds in (5), (6), *and* (8) *if and only if* $a_i = (t/n)\,[1 \leqq i \leqq n]$.

COROLLARY 2. *If* $a_i \geqq 0\,[1 \leqq i \leqq n]$ *and* $0 < \sum_{i=1}^{n} a_i \leqq 1$ *then*

(9) $$\prod_{i=0}^{n} (u_i + x_i) \geqq (n+1)^{n+1} \prod_{i=0}^{n} u_i,$$

(10) $$\sum_{i=0}^{n} \frac{x_i}{u_i} \geqq n(n+1), \qquad [u_i \neq 0]$$

(11) $$\prod_{i=0}^{n} x_i \geqq n^{n+1} \prod_{i=0}^{n} u_i.$$

Equality holds in (9), (10), *and* (11) *if and only if* $a_i = 1/(n+1)$, $[1 \leqq i \leqq n]$.

We remark that the condition

$$a_i \geqq 0[1 \leqq i \leqq n] \quad \text{and} \quad 0 < \sum_{i=1}^{n} a_i \leqq t,$$

given in the theorem, is equivalent to the condition that $P \neq A_0$ and P is a point in the simplex $A_0M_1M_2 \cdots M_n$, where $M_i = A_0 + \mathbf{a}_i t\,[1 \leqq i \leqq n]$. If $t \leqq 1$, P is in the simplex $A_0A_1 \cdots A_n$.

In the case $n = 2$, equation (3) yields the familiar result that

$$\frac{PU_0}{A_0U_0} + \frac{PU_1}{A_1U_1} + \frac{PU_2}{A_2U_2} = 1.$$

(See, for example, [2] page 162.)

The inequality (5) yields

$$A_1U_1 \cdot A_2U_2 \geqq 4\left(\frac{A_1A_2}{M_1M_2}\right)^2 (PU_1 \cdot PU_2),$$

where M_1 and M_2 are points on the rays $\mathbf{A}_0\mathbf{A}_1$ and $\mathbf{A}_0\mathbf{A}_2$, M_1M_2 is parallel to A_1A_2, and P is in the triangle $A_0M_1M_2$. Equality holds if and only if P is the midpoint of M_1M_2.

If P is in the triangle $A_0A_1A_2$, (9) yields

$$A_0U_0\cdot A_1U_1\cdot A_2U_2 \geqq 27(PU_0\cdot PU_1\cdot PU_2)$$

with equality if and only if P is the centroid.

Similar results may be obtained from the other inequalities for the case $n=2$. When $n=2$, (10) and (11) yield results obtained by Carlitz in [1].

2. Proof of the Lemma and Corollary 1. We first remark that since the vectors $\mathbf{a}_i[1\leqq i\leqq n]$ are linearly independent, it follows that $\sum_{i=1}^n a_i\neq 0$ and $a_i\neq 1\,[1\leqq i\leqq n]$.

We shall derive equation (1); (2) may be obtained in a similar manner. Let $b, c_1, c_2, \cdots, c_n$ be scalars such that

$$U_0 = P + (P - A_0)b \in A_0P \tag{12}$$

and

$$U_0 = A_1 + \sum_{i=2}^{n} (A_i - A_1)c_i \in E(0).$$

Now $P+(P-A_0)b=P+\sum_{i=1}^n \mathbf{a}_i a_i b$ and

$$A_1 + \sum_{i=2}^{n} (A_i - A_1)c_i = \left\{P - \sum_{i=1}^{n} \mathbf{a}_i a_i\right\} + \mathbf{a}_1 + \sum_{i=2}^{n} (\mathbf{a}_i - \mathbf{a}_1)c_i.$$

It follows that

$$\mathbf{a}_1\left(1 - a_1 - a_1b - \sum_{i=2}^{n} c_i\right) + \sum_{i=2}^{n} \mathbf{a}_i(c_i - a_i - a_ib) = \mathbf{0}.$$

Since $\mathbf{a}_i[1\leqq i\leqq n]$ are linearly independent, the coefficients are zero. Solving the resulting equations for b and substituting in (12) yields (1).

Immediately from the lemma we obtain

$$\frac{u_0}{u_0+x_0} = 1 - \sum_{i=1}^{n} a_i \quad \text{and} \quad \frac{u_i}{u_i+x_i} = a_i[1 \leqq i \leqq n],$$

and equations (3) and (4) follow.

3. Proof of the Theorem and Corollary 2. Since $u_i\geqq 0$ and $a_i\geqq 0$ $[1\leqq i\leqq n]$, the left side of (5) is positive. Therefore, in the derivation of (5) we may assume that $\prod_{i=1}^n u_i>0$. Since $a_i>0$ if and only if $u_i>0$, we assume that $\prod_{i=1}^n a_i>0$. Since the arithmetic mean is equal to or greater than the geometric mean,

$$\prod_{i=1}^{n} \frac{u_i + x_i}{u_i} = \prod_{i=1}^{n} \frac{1}{a_i} \geqq n^n\left(\sum_{i=1}^{n} a_i\right)^{-n} \geqq \left(\frac{n}{t}\right)^n$$

with equality if and only if $a_i=t/n\,[1\leqq i\leqq n]$.

The inequality (6) is derived similarly. If $u_i\neq 0\,[1\leqq i\leqq n]$,

$$\sum_{i=1}^{n} \frac{x_i}{u_i} = \sum_{i=1}^{n} \frac{1-a_i}{a_i} = \left(\prod_{i=1}^{n} a_i\right)^{-1} \sum_{j=1}^{n} \left(\prod_{1 \le i \le n, i \ne j} a_i\right) - n$$
$$\geqq n\left(\prod_{i=1}^{n} a_i\right)^{-1/n} - n \geqq n^2\left(\sum_{i=1}^{n} a_i\right)^{-1} - n \geqq \frac{n^2}{t} - n.$$

To derive (7) let $I = \{i: 1 \leqq i \leqq n,\ u_i \neq 0\}$. Now $I \neq \emptyset$ because $u_i = 0$ if and only if $a_i = 0$, and $\sum_{i=1}^{n} a_i > 0$. Let

$$z = \min_{i \in I} \frac{x_i}{u_i} = \min_{i \in I} \frac{1-a_i}{a_i} \geqq \frac{1-t}{t}\cdot$$

Since $u_i \geqq 0$, $x_i \geqq u_i z$ and $\sum_{i \in I} x_i \geqq z \sum_{i=1}^{n} u_i$. If $I = \{i: 1 \leqq i \leqq n\}$ then

$$\sum_{i=1}^{n} x_i \geqq z \sum_{i=1}^{n} u_i > \frac{1-t}{t} \sum_{i=1}^{n} u_i.$$

If $u_j = 0$ then $x_j > 0$ and

$$\sum_{i=1}^{n} x_i > \sum_{i \in I} x_i \geqq \frac{1-t}{t} \sum_{i=1}^{n} u_i.$$

This proves (7).

The left side of (8) is nonnegative. Therefore in the derivation of (8) we may assume that $\prod_{i=1}^{n} u_i > 0$. That is, we may assume that $a_i > 0\,[1 \leqq i \leqq n]$.

Now (8) is satisfied when $n = 2$. For if a_1 and a_2 are positive scalars such that $0 < a_1 + a_2 \leqq t \leqq 1$, then

$$\frac{x_1 x_2}{u_1 u_2} = \frac{(1-a_1)(1-a_2)}{a_1 a_2} \geqq \left(\frac{2}{t} - 1\right)^2$$

with equality if and only if $a_1 = a_2 = t/2$.

We assume that (8) is satisfied for $n = k$, and suppose that $a_i[1 \leqq i \leqq k+1]$ are positive scalars such that $0 < \sum_{i=1}^{k+1} a_i = s \leqq t \leqq 1$. Let $\sum_{i=1}^{k} a_i = z < t$. Then $a_{k+1} = s - z$ and

$$\prod_{i=1}^{k+1} \frac{x_i}{u_i} = \left\{\prod_{i=1}^{k} \frac{(1-a_i)}{a_i}\right\} \cdot \frac{1-a_{k+1}}{a_{k+1}}$$
$$\geqq \left(\frac{k}{z} - 1\right)^k \left(\frac{1-s+z}{s-z}\right) \geqq \left(\frac{k}{z} - 1\right)^k \left(\frac{1-t+z}{t-z}\right)$$

with equality if and only if $a_i = (z/k)\,[1 \leqq i \leqq k]$ and $s = t$.

The proof is completed by showing that for $0 < z < t \leqq 1$,

$$\left(\frac{k}{z} - 1\right)^k \left(\frac{1-t+z}{t-z}\right) - \left(\frac{k+1}{t} - 1\right)^{k+1} \geqq 0$$

with equality if and only if $z = kt/(k+1)$.

We remark that (8) is not valid for $t>1$.

To derive the results in Corollary 2 we note that if $0<\sum_{i=1}^{n} a_i=t<1$ then from (5), (6), (8), and (3), we obtain:

$$\prod_{i=0}^{n}(u_i+x_i) \geqq \left(\frac{n}{t}\right)^n\left(\frac{1}{1-t}\right)\prod_{i=0}^{n} u_i = f_n(t)\prod_{i=0}^{n} u_i$$

$$\sum_{i=0}^{n}\frac{x_i}{u_i} \geqq n\left(\frac{n}{t}-1\right)+\frac{t}{1-t} = g_n(t) \quad [u_i \neq 0]$$

$$\prod_{i=0}^{n} x_i \geqq \left(\frac{n}{t}-1\right)^n\left(\frac{t}{1-t}\right)\prod_{i=0}^{n} u_i = h_n(t)\prod_{i=0}^{n} u_i.$$

The proof of Corollary 2 is completed by noting that for $t\in(0, 1)$,

$$\min f_n(t) = f_n\left(\frac{n}{n+1}\right) = (n+1)^{n+1},$$

$$\min g_n(t) = g_n\left(\frac{n}{n+1}\right) = n(n+1),$$

$$\min h_n(t) = h_n\left(\frac{n}{n+1}\right) = n^{n+1}.$$

This work was supported by National Science Foundation Grant G23554.

References

1. L. Carlitz, Some inequalities for a triangle, this MONTHLY, 71(1964) 881–885.
2. N. A. Court, College Geometry, Barnes and Noble, New York, 1952.
3. University of Illinois Committee on School Mathematics (UICSM), A vector approach to Euclidean geometry (in preparation).

AN EXTENSION OF A THEOREM OF RAMANUJAN

JOSEPH ARKIN, Suffern, New York

Introduction. Let $f(x)=\prod_{n=1}^{\infty}(1-x^n)$. Following Newman [1] we write

$$(f(x))^k = \sum_{n=0}^{\infty} p_k(n)x^n.$$

Ramanujan [2] proved in 1919 that $p_6(n)\equiv 0 \pmod{49}$ for $n\equiv -2 \pmod 7$. In the following theorem we have replaced 7 by any prime $P\equiv 3 \pmod 4$.

THEOREM I. *Let $P=4m+3$ be a prime then*

(1) $$p_6(n) \equiv 0 \pmod{P^2} \qquad \text{for } n \equiv -1-m \pmod{P}.$$

Proof. From Jacobi's well-known identity [3]

$$(f(x))^3 = \sum_{s=0}^{\infty} (-1)^s(2s+1)x^{s(s+1)/2}$$

we have

$$(f(x))^6 = \sum_{s=0}^{\infty}\sum_{t=0}^{\infty} (-1)^{s+t}(2s+1)(2t+1)x^{s(s+1)/2+t(t+1)/2}.$$

Now let $n=\mu P-m-1$; then $p_6(n)$ is the coefficient of $x^{\mu P}$ in the expansion of $x^{m+1}(f(x))^6$. Hence

$$p_6(n) = \sum\sum (-1)^{s+t}(2s+1)(2t+1), \tag{1.1}$$

where the sum extends over all nonnegative s, t satisfying

$$s(s+1)/2 + t(t+1)/2 + m + 1 = \mu P.$$

The condition under the summation can be written, after multiplication by 8,

$$(2s+1)^2 + (2t+1)^2 = (8\mu-2)P. \tag{1.2}$$

Since -1 is not a quadratic residue of P, P can divide the left side of (1.2) only by dividing each term. Now (1) follows from (1.1).

In a previous paper [4] I proved the following result. Put

$$\left(\sum_{n=0}^{\infty} c_n x^n\right)^k = \sum_{n=0}^{\infty} c_n^{(k)} x^n, \tag{2}$$

where the c_n are integers and k is an integer $\geqq 1$. We define $c_n^{(k)}$ by means of (2) for all integral k. (The series considered are purely formal power series, and the question of their convergence is irrelevant.) We then have:

$$c_n^{(k)} \equiv 0 \pmod{k/(n,k)} \qquad n = 1, 2, 3, \cdots, \tag{2.1}$$

and, when $c_0=1$,

$$c_n^{(-k)} \equiv 0 \pmod{k/(n,k)}. \tag{2.2}$$

The congruences in (2) then led this author to the following supplementary results [5].

(3) If for a prime P and for some integers t and $r (r\geqq 1)$, the congruence $c_v^{(d)}\equiv 0 \pmod{P^{r-s}}$, $(v=mP+t)$, holds for $m=0, 1, 2, \cdots$ and for $0\leqq s\leqq r$, then

$$c_v^{(w+d)} \equiv 0 \pmod{P^{r-s}} \qquad (w = P^r a)$$

for any integer a.

(3.1) If $c_0=1$, and if $c_v^{(\pm d)}\equiv 0 \pmod{P^{r-s}}$, P being a prime, and t is any

integer for which the congruence holds, then $c_{s}^{(\pm w\pm d)}\equiv 0 \pmod{P^{r-s}}$. Now let $\sum_{n=0}^{\infty} p_6(n)x^n=\sum_{n=0}^{\infty} c_n^{(6)}x^n$, so that by (1, with $P=7$) we have

$$c_{7m+5}^{(6)} \equiv 0 \pmod 7 \qquad m = 0, 1, 2, \cdots.$$

Combining this result with (3.1, with $-w=7$, $d=6$) leads to

$$c_{7m+5}^{(-1)} = p_{-1}(7m+5) \equiv 0 \pmod 7$$

which is a well-known congruence of Ramanujan.

In conclusion we discuss some consequences of (2) and (3). Let

$$\left(\sum_{n=0}^{\infty} \sigma(2n+1)x^n\right)^{-z} = \sum_{n=0}^{\infty} A_n x^n,$$

where $\sigma(2n+1)$ equals the sum of the divisors of $2n+1$.

THEOREM II. *If $P=4z+1$ is a prime number $z=1, 2, 3, \cdots$, and t is any integer for which $Pm+t\neq v(v+1)/2$, $((P, t)=1, m=0, 1, 2, \cdots, v=0, 1, 2, \cdots)$, then*

$$A_{Pm+t} \equiv 0 \pmod P. \tag{4}$$

Proof. Put

$$\left(\sum_{v=0}^{\infty} x^{v(v+1)/2}\right)^{P} = \sum_{n=0}^{\infty} c_n^{(P)} x^n \qquad (\text{here } c_0 = 1). \tag{4.1}$$

Since $c_0=1$ and $(P, t)=1$, combining (4.1) with (2.2, with $-k=P$) leads to

$$c_{Pm+t}^{(-P)} \equiv 0 \pmod P. \tag{4.2}$$

Now, combining $Pm+t\neq v(v+1)/2$ with (4.1), we have $0=c_{Pm+t}^{(1)}\equiv 0 \pmod P$, and this result with (4.2) and (3.1, with $s=0$, $r=1$, $-w=P=4z+1$, $d=1$) leads to

$$c_{Pm+t}^{(-4z)} \equiv 0 \pmod P. \tag{4.3}$$

A well-known identity of Jacobi (see [6]) is

$$\sum_{n=0}^{\infty} c_n^{(-4z)} x^n = \left(\sum_{v=0}^{\infty} x^{v(v+1)/2}\right)^{-4z} = \left(\sum_{n=0}^{\infty} \sigma(2n+1)x^n\right)^{-z} \tag{4.4}$$

$(z=1, 2, 3, \cdots)$, and (4) now follows from (4.3) and (4.4). The following result follows easily.

COROLLARY. *If $P=4z-1$ is a prime number, then*

$$c_{Pm+t}^{(4z)} \equiv 0 \pmod P.$$

Acknowledgments. The author is very grateful to Professor R. Pollack for his encouragement and guidance and to the referee for his valuable suggestions.

References

1. M. Newman, Some theorems about $p_r(n)$, Canad. J. Math., 9(1957)549–552.
2. S. Ramanujan, Proc. Cambridge Phil. Soc., 19 (1919) 207–210.
3. C. G. J. Jacobi, Fundamenta Nova, 1829, 66, (7); Werke, 1, 237.
4. J. Arkin, Congruences for the coefficients of the kth power of a power series, this MONTHLY. 71(1964)899–900.
5. ———, Congruences for the coefficients of the kth power of a power series (II), this MONTHLY, 72(1965)390–393.
6. C. G. J. Jacobi, Dickson's Theory of Numbers, vol. 1, p. 281.

MATHEMATICAL NOTES

EDITED BY J. H. CURTISS, University of Miami

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

AN INEQUALITY FOR THE DERIVATIVES OF NONNEGATIVE POLYNOMIALS

L. F. SHAMPINE, Sandia Laboratory, Albuquerque, New Mexico

We shall prove an analog for nonnegative polynomials of the following pretty inequality of Turán [1]:

THEOREM. *Let $P_n(x)$ be a polynomial of degree $n \geqq 2$ and $L_j(x)$ be the Laguerre polynomials. Then*

$$\int_0^\infty [P_n'(x)]^2 e^{-x}dx \leqq \tfrac{1}{4}\sin^{-2}\frac{\pi}{4n+2}\int_0^\infty [P_n(x)]^2 e^{-x}dx$$

with equality holding if and only if P_n is a multiple of

$$\sum_{j=1}^{n} L_j(x)\sin\frac{j\pi}{2n+1}\cdot$$

Primes denote differentiation and the Laguerre polynomials are explicitly

$$L_j(x) = \sum_{k=0}^{j}(-1)^k\frac{1}{k!}\binom{j}{k}x^k \qquad j = 0, 1, \cdots.$$

Polynomials P_n such that $P_n(x) \geqq 0$ for $x \geqq 0$ are characterized [2, p. 185] by

$$P_n(x) = \left|\sum_{0}^{[n/2]} u_j L_j(x)\right|^2 + x\left|\sum_{0}^{[(n-1)/2]} v_j r_j(x)\right|^2, \tag{1}$$

where the L_j and r_j are orthonormal with respect to e^{-x} and xe^{-x}. The u_j and v_j may be complex numbers.

THEOREM. *Let $P_n(x)$ be a polynomial of degree $n \geqq 2$ and suppose $P_n(x) \geqq 0$ for $x \geqq 0$. Then*

$$-\left(\left[\frac{n}{2}\right]-1\right)\int_0^\infty P_n(x)e^{-x}dx \leqq \int_0^\infty P_n'(x)e^{-x}dx \leqq \int_0^\infty P_n(x)e^{-x}dx.$$

Equality on the left is obtained if and only if P_n is a nonnegative multiple of

$$\left(\sum_{j=1}^{[n/2]} L_j(x)\right)^2.$$

Equality on the right is obtained if and only if P_n, when expressed in the form (1), *has*

$$u_0 = 0, \qquad \sum_{j=1}^{[n/2]} u_j = 0.$$

With the requirement $P_n(x) \geqq 0$ for $x \geqq 0$ we ask for the extrema of

$$\int_0^\infty P_n'(x)e^{-x}dx$$

subject to

$$\int_0^\infty P_n(x)e^{-x}dx = 1.$$

Using the characterization, we see that

$$\int_0^\infty P_n e^{-x}dx = \sum_0^{[n/2]} \left|u_j\right|^2 + \sum_0^{[(n-1)/2]} \left|v_j\right|^2 \tag{2}$$

and

$$P_n' = \sum u_i\bar{u}_j(L_iL_j)' + \sum v_i\bar{v}_j(xr_ir_j)'.$$

An integration by parts and the fact that $L_j(0) = 1$ for all j show that

$$\int_0^\infty P_n' e^{-x}dx = \sum_{i,j=1}^{[n/2]} u_i\bar{u}_j(\delta_{ij} - 1) + \sum_0^{[(n-1)/2]} |v_j|^2. \tag{3}$$

The extremum of the quadratic form is obtained from those of the separate forms which we now determine. All the eigenvalues of the v-form are obviously $+1$. There is a zero eigenvalue associated with u_0. Let J be the $[n/2]\times[n/2]$ matrix all of whose entries are $+1$. The matrix associated with u-form (excluding u_0) is $-M$ with $M = J - I$.

Since $J^2=[n/2]J$ we readily find the minimal polynomial of M to be

$$\lambda^2-\left(\left[\frac{n}{2}\right]-2\right)\lambda-\left(\left[\frac{n}{2}\right]-1\right).$$

The roots are -1 and $[n/2]-1$. M is a nonnegative irreducible matrix. It is known that the largest eigenvalue in modulus of such matrices is simple [**3**, p. 30]. The eigenvector corresponding to this value is obviously the vector all whose entries are 1.

The only negative eigenvalue of the form is $-([n/2]-1)$, so that

$$-\left(\left[\frac{n}{2}\right]-1\right)\left[\sum_0^{[n/2]}|u_i|^2+\sum_0^{[(n-1)/2]}|v_i|^2\right]\leqq\int_0^\infty P_n'(x)e^{-x}dx$$

which (2) shows to be the left side of the desired inequality. We saw equality could occur if and only if the vector is a multiple of

$$u_0=0,\qquad u_i=1,\qquad i=1,\cdots,\left[\frac{n}{2}\right]$$
$$v_i=0,\qquad i=0,1,\cdots,\left[\frac{n-1}{2}\right].$$

From (1) we see that this means equality holds if and only if P_n is a nonnegative multiple of

$$\left(\sum_{i=1}^{[n/2]}L_i(x)\right)^2.$$

The remaining eigenvalues are 0 and the multiple eigenvalue $+1$. As before, this statement is the right side of the inequality. Equality is attained if and only if the vector lies in the invariant subspace corresponding to $+1$. This is equivalent to the vector being orthogonal to the eigenvectors corresponding to 0 and $-([n/2]-1)$, i.e.

$$u_0=0,\qquad \sum_{i=1}^{[n/2]}u_i=0.$$

This work was supported by the U. S. Atomic Energy Commission. Reproduction in whole or in part is permitted for any purpose of the U. S. Government.

References

1. P. Turán, Remark on a theorem of Erhard Schmidt, Mathematica (Cluj), 2 (25), 2 (1960), 373–78.

2. G. Szegö, Orthogonal polynomials, Amer. Math. Soc. Colloq. Publ., 23 (1939).

3. R. S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, N. J., 1962.

ON THE FUNCTIONAL EQUATION $f^n+g^n=h^n$

FRED GROSS, U. S. Naval Research Laboratory, Washington, D. C.

I. Introduction. The purpose of this paper is to study the solutions of functional equations of the form

$$(f(z))^n + (g(z))^n = (h(z))^n \tag{1}$$

where n is an integer greater than 1.

Some of our more interesting results are stated below.

I. Let n be an integer greater than 2. If $p(z) \not\equiv 0$ is a polynomial of degree $k \leqq n-2$ then the functional equation

$$(f(z))^n + (g(z))^n = p(z) \tag{2}$$

does not have any nonconstant entire solutions.

II. The functional equation (2) with $p(z) \equiv 1$ and $n=2$ has the entire solutions $f(z) = \cos(\eta(z))$ and $g(z) = \sin(\eta(z))$, where $\eta(z)$ is any entire function. No other solutions exist.

III. If $p(z)$ is any polynomial, then the functional equation

$$(f(z))^n + (g(z))^n = (p(z))^n, \ (n > 2) \tag{3}$$

has no entire solutions other than constants.

A number of related problems will also be discussed.

We shall need a particular consequence of the following theorem due to Borel [1]. By the expression $F(z)$ grows more slowly (faster) than $\nu(r)$ we mean that the maximum modulus $M(r, F)$ of $F(z)$ is eventually less (greater) than $\nu(r)$.

THEOREM (BOREL). *Let k be a positive integer and $\mu(r)$ be a positive increasing function. Let $G_i(z)$ and $H_i(z)$, $i=1, 2, \cdots, k$, be entire functions such that $G_i(z)$ grows more slowly than $e^{\mu(r)}$ and $H_i(z)-H_j(z)$ grows faster than $[\mu(r)]^2$ for $i \neq j$. If $\sum_{i=1}^{k} G_i(z)e^{H_i(z)}=0$ then $G_i(z)=0$ for $i=1, 2, \cdots, k$.*

An immediate consequence is the following

LEMMA 1. *Let $p_i(z)$ be polynomials and $\phi_i(z)$ be entire functions for $i=1, 2, 3$. If $p_1(z)e^{\phi_1(z)}+p_2(z)e^{\phi_2(z)}=p_3(z)$ then $\phi_1(z)$ and $\phi_2(z)$ are both constants.*

THEOREM 1. *Let $p(z, w)$ be an entire function in two complex variables of the form $(a_1z+b_1w)(a_2z+b_2w)(a_3z+b_3w)Q(z, w)$, where $Q(z, w)$ is an entire function in z and w. Suppose furthermore that*

$$D_{ij} \equiv \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix} \neq 0 \qquad \text{for } i \neq j \ (i, j = 1, 2, 3).$$

If $f(z)$ and $g(z)$ are any two entire functions such that $p(f(z), g(z))$ has at most a finite number of zeros, then $f(z)/g(z)$ must be a rational function of z.

Proof. We may assume that $a_1 \neq 0$. We have

$$a_1 f(z) + b_1 g(z) = p_1(z) e^{\phi_1(z)} \tag{4}$$

$$a_2 f(z) + b_2 g(z) = p_2(z) e^{\phi_2(z)} \tag{5}$$

and

$$a_3 f(z) + b_3 g(z) = p_3(z) e^{\phi_3(z)} \tag{6}$$

where $p_i(z)$ are polynomials and $\phi_i(z)$ are entire functions for $i=1, 2, 3$. Eliminating $f(z)$ from equations (4), (5) and (6) respectively and solving for $g(z)$ we get

$$\frac{a_2 p_1(z) e^{\phi_1(z)} - a_1 p_2(z) e^{\phi_2(z)}}{a_2 b_1 - a_1 b_2} = g(z) = \frac{a_3 p_1(z) e^{\phi_1(z)} - a_1 p_3(z) e^{\phi_3(z)}}{a_3 b_1 - a_1 b_3}.$$

By Lemma 1 it follows that $\phi_2(z) - \phi_3(z)$ is equal to a constant. Hence (5) and (6) yield

$$\left[\frac{a_2}{p_2(z)} - \frac{c a_3}{p_3(z)}\right] f(z) = \left[\frac{c b_3}{p_3(z)} - \frac{b_2}{p_2(z)}\right] g(z),$$

where c is a nonzero constant. Since $D_{ij} \neq 0$ if $i \neq j$, by hypothesis, the coefficients of $f(z)$ and $g(z)$ do not vanish identically and hence $f(z)/g(z)$ must be a rational function.

THEOREM 2. *Let $f(z)$ and $g(z)$ be any two entire functions such that $f(z)/g(z)$ is not a rational function of z; then the function $(f(z))^n + (g(z))^n$, $(n > 2)$, must have infinitely many zeros.*

Proof. Let ξ_1, ξ_2 and ξ_3 be three distinct nth roots of -1. We have

$$(f(z))^n + (g(z))^n = [f(z) - \xi_1 g(z)][f(z) - \xi_2 g(z)][f(z) - \xi_3 g(z)]Q(z),$$

where $Q(z)$ is an entire function. Since

$$\begin{vmatrix} 1 & 1 \\ \xi_i & \xi_j \end{vmatrix} \neq 0 \qquad \text{for } i \neq j$$

our result follows from Theorem 1.

THEOREM 3. *Let n be an integer greater than 2. If $p(z) \neq 0$ is a polynomial of degree $k \leq n-2$ then the functional equation* (2) *does not have any nonconstant entire solutions.*

Proof. If $f(z)$ and $g(z)$ satisfy (2), then it follows from Theorem 2 that $f(z)$ and $g(z)$ must be polynomials. This is clear, since $f(z)/g(z)$ must be rational and so $f(z) = r(z)g(z)$, where $r(z)$ is rational. Thus $(1 + r(z))^n (g(z))^n = p(z)$. If $g(z)$ were

transcendental then the left side of this equation would grow faster than the right side.

Factoring the two sides of (2) we get

$$[f(z)+\xi_1 g(z)]\cdots[f(z)+\xi_n g(z)] = c(z-a_1)\cdots(z-a_k),$$

where c is a constant and the ξ_i are roots of unity. Since $k \leqq n-2$, two factors on the left side of the above equation must reduce to nonzero constants and consequently $g(z)$ must be constant.

In a similar manner one sees that $f(z)$ must also be a constant.

THEOREM 4. *The functional equation* $(f(z))^n+(g(z))^n=1$ *does not have any nonconstant entire solutions for* $n>2$, *while for* $n=2$ *all entire solutions* $f(z)$ *and* $g(z)$ *are of the form* $f(z)=\cos(\eta(z))$ *and* $g(z)=\sin(\eta(z))$, *where* $\eta(z)$ *is an entire function.*

Proof. The first part of the theorem follows immediately from Theorem 3. We prove the second part.

$$(f(z))^2+(g(z))^2 = 1$$

implies that $[f(z)+ig(z)][f(z)-ig(z)]=1$. Hence we must have

$$f(z)-ig(z) = e^{\phi(z)} \tag{7}$$

$$f(z)+ig(z) = e^{-\phi(z)}, \tag{8}$$

where $\phi(z)$ is an entire function. Solving the simultaneous equations (7) and (8) for $f(z)$ and $g(z)$ we get

$$f(z) = \frac{e^{\phi(z)}+e^{-\phi(z)}}{2} \quad\text{and}\quad g(z) = \frac{e^{-\phi(z)}-e^{\phi(z)}}{2i}$$

and our theorem follows.

It follows from Theorem 2 as in the proof of Theorem 3 that for any polynomial $p(z)$, the functional equation (3) has no entire solutions other than polynomials. But it is known (see Shanks [4]) that nonconstant polynomial solutions do not exist either. Thus we have

THEOREM 5. *If* $p(z)$ *is any polynomial, then the functional equation* (3) *has no entire solutions other than constants.*

REMARK. For $n=2$ it is well known that $(f(z))^2+(g(z))^2=1$ does have rational functions as solutions [4].

Let us now consider the functional equation (1) with $n>2$. It follows from Theorem 5 that if $f(z)$, $g(z)$ and $h(z)$ satisfy equation (1) then each of them must have infinitely many zeros.

The problem of determining whether such solutions actually exist seems to be quite difficult, but we have:

THEOREM 6. *There exist functions $f(z)$, $g(z)$ and $h(z)$, analytic in the open unit disc which satisfy equation* (1) *with* $n=4$.

Proof. Let $f(q)=\theta_2(q^4)$, $g(z)=\theta_4(q^4)$ and $h(q)=\theta_3(q^4)$, with $\theta_i(q)=\theta_i(0, q)$, where $\theta_i(z, q)$ are the Jacobi theta functions for $i=2, 3, 4$.

$$\theta_2(q) = 2\sum_{n=0}^{\infty} q^{(1/4)(2n+1)^2}, \quad \theta_3(q) = 1 + 2\sum_{n=1}^{\infty} q^{n^2}, \quad \theta_4(q) = 1 + 2\sum_{n=1}^{\infty} (-1)^n q^{n^2}.$$

It is well known (see Hille [2]) that $\theta_2(q)$, $\theta_3(q)$ and $\theta_4(q)$ satisfy the equation $(\theta_2(q))^4+(\theta_4(q))^4=(\theta_3(q))^4$. Thus

$$(f(q))^4 + (g(q))^4 = (h(q))^4.$$

Since $f(q)$, $g(q)$ and $h(q)$ are clearly analytic in $|q|<1$ our theorem follows.

It is interesting to note that the more general functional equation

$$(f_1(z))^n + (f_2(z))^n + \cdots + (f_k(z))^n = 1 \tag{9}$$

may have entire solutions. For $n=k=3$ (Lehmer [3]) showed that (9) has infinitely many solutions, one of which is given by

$$f_1(z) = 9z^4, \quad f_2(z) = -9z^4+3z \quad \text{and} \quad f_3(z) = -9z^3+1.$$

Entire solutions of (9) also exist when $n>k$ as illustrated by the example

$$[2^{1/4}(\sin^2 z - \cos^2 z + i\sin z\cos z)]^4 + [(-1)^{1/4}(2i\sin z\cos z + \sin^2 z)]^4 + [(-1)^{1/4}(2i\sin z\cos z - \cos^2 z)]^4 = 1.$$

In a subsequent paper the author will prove that (1) has no meromorphic solutions for $n>3$ and will discuss the solutions of (1) for $n=3$. The proofs are simple, but more advanced methods must be used than are employed in this paper.

Subsequent to the completion of this paper, a paper containing a number of similar results appeared (see—Elementary proof of a theorem of P. Montel, on entire functions by A. V. Jategaonkar, J. London Math. Soc., 40 (1965) 166–170). It has also been brought to the author's attention by Prof. Olga T. Todd, that Ganapathy Iyer had done some work on (1). In fact he also proves the second part of Theorem 4 (see his paper "On certain functional equations," J. Indian Math. Soc., 3 (1939) 312–315).

The author was formerly with the National Bureau of Standards.

References

1. E. Borel, Sur les zéros des fonctions entières, Acta Mathematica, 20 (1897) 357–396.
2. E. Hille, Analytic Function Theory, vol. II, Ginn, Boston, Mass., 1962, p. 159.
3. D. H. Lehmer, On the Diophantine Equation $x^3+y^3+z^3=1$, J. London Math. Soc., 31 (1956) 275–280.
4. D. Shanks, Solved and Unsolved Problems in Number Theory, vol. I, Spartan Books, Washington, D. C., 1962, pp. 145–146.

ON THE CLASS OF YOUNG'S CONTINUOUS FUNCTIONS

B. S. YADAV, Sardar Patel University, Vallabh Vidyanagar, India

1. Let the functions f, g and h be each L-integrable in $(0, 2\pi)$ and periodic outside with period 2π. We say that h is the Fourier faltung (composition or convolution) of f and g if

$$h(x) = \frac{1}{2\pi}\int_0^{2\pi} f(x+t)g(t)dt, \qquad [7; \text{p. } 36]. \tag{1}$$

Moreover, a function h is said to be a Young's continuous function if there exist two functions f and g, each of the Lebesgue class L_2, such that h is the Fourier faltung of these functions [6]. Although there is no obvious criterion to verify directly whether a given function h is a Young's continuous function, it is known that a necessary and sufficient condition for a function to be a Young's continuous function is that its Fourier series be absolutely convergent ([1], [3]). Min-Teh Cheng [2] has obtained Young's continuous functions by imposing more on f and less on g as follows:

THEOREM 1. *If* $f \in \text{lip}(\alpha, p)$ *and* $g \in \text{lip}(\frac{1}{2}p, q)$, *where* $0 < \alpha \leqq 1$, $1 < p \leqq 2$, $\alpha p > \frac{1}{2}$ *and* $q > 1$, *then the function* h *given by* (1) *is a Young's continuous function.*

The object of this note is also to obtain Young's continuous functions with the help of the following theorem proved first by the author in [5].

THEOREM 2. *Let* $0 < \alpha \leqq 1$, $1 < p \leqq 2$, *and* $t > 0$. *If*

$$f \sim \sum_{-\infty}^{\infty} c_n e^{inx} \tag{2}$$

and

$$\int_0^{2\pi} |f(x+t) - f(x)|^p dx = O(t^{\delta}(\log t^{-1})^{-1-\alpha p}), \qquad \text{as } t \to 0,$$

where $\delta = 1 + p(1-\beta)/\beta$, *then*

$$\sum_{-\infty}^{\infty}{}' \; |c_n|^{\beta}(\log|n|)^T < \infty \tag{3}$$

for all $\beta > p(T+1)/(1+\alpha p)$; *but not necessarily for* $\beta = p(T+1)/(1+\alpha p)$.

($\sum'$ denotes a summation for n in which the term corresponding to $n = 0$ is omitted.)

We shall prove the following

THEOREM 3. *Let* $0 < \alpha \leqq 1$, $1 < p \leqq 2$ *and* $t > 0$. *If*

$$\int_0^{2\pi} |(f(x+t) - f(x)|^p dx = O(t^{2-p}(\log t^{-1})^{-1-\alpha p}), \qquad \text{as } t \to 0, \tag{4}$$

and $g \in \text{lip}(1/q, q)$, *where* $p(1-\alpha)<1$ *and* q *is given by* $p^{-1}+q^{-1}=1$, *then the function* h *in* (1) *is a Young's continuous function.*

The case $p=2$ in this theorem is trivial.

2. In order to prove our theorem, we shall need the following three lemmas:

LEMMA 1. *If* $f \in \text{lip}(\alpha, p)$, *then* $f \in L_p$.

LEMMA 2. *If* $f \in \text{lip}(\alpha, p)$, $\alpha>0$ *and* $p>1$, *then*

$$c_n = O(|n|^{\alpha}).$$

LEMMA 3. *If* f *satisfies the condition* (4), *then* $\sum_{-\infty}^{\prime\infty} |c_n(\log|n|)^T|^p < \infty$, *for* $T<\alpha$.

The Lemmas 1 and 2 are obtained by Hardy and Littlewood [4] and the Lemma 3 can be obtained as a corollary of Theorem 1 by replacing β by p and T by Tp in (3).

3. *Proof of Theorem* 3. We first observe that the integral

$$\int_0^{2\pi} f(x+t)g(t)dt$$

does exist under the hypothesis of the theorem. For, it follows from the condition (4) that $f \in \text{lip}((2-p)/p, p)$; and hence from Lemma 1 that $f \in L_p$ and $g \in L_q$. Therefore, by Hölder's inequality, we have

$$\left|\int_0^{2\pi} f(x+t)g(t)dt\right| \leqq \left\{\int_0^{2\pi} |f(x+t)|^p dt\right\}^{1/p} \left\{\int_0^{2\pi} |g(t)|^q dt\right\}^{1/q} < \infty.$$

To prove the theorem it will suffice to prove that the Fourier series of h is absolutely convergent. Let $g \sim \sum_{-\infty}^{\infty} d_n e^{inx}$. Then we know that $h \sim \sum_{-\infty}^{\infty} c_n d_{-n} e^{inx}$. Now

$$\sum_{-N}^{N}{}^{*} |c_n d_{-n}| \leqq \left\{\sum_{-N}^{N}{}' |c_n(\log|n|)^T|^p\right\}^{1/p} \left\{\sum_{-N}^{N}{}' |d_{-n}(\log|n|)^{-T}|^q\right\}^{1/q}, \tag{5}$$

where in $\sum^*$ the terms with $|n| \leqq 1$ are omitted, and where T is chosen so that $T<\alpha$ and $Tq>1$.

This choice of T is possible because of the hypothesis $p(1-\alpha)<1$. It follows from Lemma 3 that the first sum on the right of (5) remains bounded, as $N \to \infty$, under the condition (4). Again it follows from Lemma 2 that $d_{-n}=O(|n|^{1/q})$. Therefore

$$\sum_{-N}^{N}{}' |d_{-n}(\log|n|)^{-T}|^q = O\left(\sum_{-N}^{N}{}' \frac{1}{n(\log|n|)^{Tq}}\right).$$

Since $Tq>1$, the second sum on the right of (5) will also remain bounded, as

$N \to \infty$. Therefore

$$\sum_{-\infty}^{\infty} |c_n d_{-n}| < \infty,$$

and the proof is complete.

I wish to thank Professor U. N. Singh for his help in the preparation of this note.

References

1. K. K. Chen, On the class of functions with absolutely convergent Fourier series, Proc. Imp. Acad., 4 (1928) 348–351.

2. Min-Teh Cheng, The absolute convergence of Fourier series, Duke Math. J., 9 (1942) 803–810.

3. G. H. Hardy and J. E. Littlewood, Notes on the theory of series (IX): On the absolute convergence of Fourier series, J. London Math. Soc., 3 (1928) 250–253.

4. ———, A convergence criterion for Fourier series, Math. Z., 28 (1928) 612–634.

5. B. S. Yadav, On a theorem of Titchmarsh, Publ. Inst. Math., Beograd, Nouvelle série, tome 2 (16) (1962) 87–92.

6. W. H. Young, On a class of parametric integrals and their application in the theory of Fourier series, Proc. Roy. Soc., A, 25 (1911) 401.

7. A. Zygmund, Trigonometric series, vol. 1, Cambridge, 1959.

ON RELATIVELY PRIME SEQUENCES

M. V. SUBBARAO, University of Alberta, Edmonton

1. A sequence of integers $\{a_n\}$ $(n \geqq 1)$ will be called relatively prime if $a_n \neq 0$ except for at most one value of n and $(a_m, a_n) = 1$ for all m and n with $m \neq n$. We recall that, if a is a nonzero integer, we define $(a, 0) = |a|$. In this note we consider the following question: Suppose that $\phi(x)$ is a polynomial in x with integral coefficients and $\phi_0(x) = x$, $\phi_{k+1}(x) = \phi(\phi_k(x))$ $(k \geqq 0)$. Suppose also that for all integers x and all $k \geqq 0$ for which $\phi_k(x) \neq 0$ we have $\phi_k(x) \mid \phi_{k+1}(x)$. We note that this holds if and only if $\phi(0) = 0$, which is therefore assumed throughout. Under what conditions on $\phi(x)$ and for what integral values of x is the sequence of integers $\{f_n(x)\}$ $(n > 0)$ relatively prime, where $f_n(x) = \phi_n(x)/\phi_{n-1}(x)$ $(n = 1, 2, \cdots)$? (For the case $\phi_{n-1}(x) = 0$, see below.) This question arises naturally in view of the known result that the sequence $\{(a^{p^n} - 1)/(a^{p^{n-1}} - 1)\}$ $(n \geqq 1$, p prime, $(a-1, p) = 1)$ is relatively prime, and we can write $a^{p^n} - 1 = \phi_n(a-1)$ with $\phi(x) = (x+1)^p - 1$.

A similar problem for the sequence $\{\phi_n(x)\}$ of iterations of $\phi(x)$ was considered by R. Bellman [1] in trying to generalize the well-known result that the sequence of Fermat numbers $\{F_n\} = \{2^{2^n} + 1\}$ is relatively prime and observing the fact that F_n can be written as $F_n = \phi_n(3)$, where $\phi(x) = (x-1)^2 + 1$. In the sequel $\{\phi_n(x)\}$ and $\{f_n(x)\}$ are sequences as already defined and we write $f(x)$ for $f_1(x)$.

If, for some $n \geqq 0$, $\phi_n(x) = 0$, then $\phi_m(x) = 0$ for all $m \geqq n$. In this case we define

The author is thankful to the referee especially for calling attention to the case in which $\phi_n(x)$ vanishes for some n.

References

1. Richard Bellman, A note on relatively prime sequences, Bull. Amer. Math. Soc., 53 (1947) 778–779.
2. J. Lambek and L. Moser, On relatively prime sequences, Math. Gazette, 41 (1957) 287–288.

THE MASSERA—SCHAFFER EQUALITY

L. M. KELLY, Michigan State University and Cambridge University

J. L. Massera and J. J. Schaffer [Annals of Math., 67 (1958) p. 538] have shown that for any two vectors x and y of a normed linear space

$$\| \|y\|x - \|x\|y\| \max(\|x\|, \|y\|) \leqq 2\|x\| \cdot \|y\| \cdot \|x - y\|. \tag{1}$$

Kirk and Smiley [this MONTHLY, 71 (1964) p. 891] have expressed interest in the conditions under which equality holds.

THEOREM. *Two distinct nonzero vectors x and y of a complex normed linear space satisfy the equality in* (1) *iff x and y span an l_2^1 in the underlying real vector space with*

$$\pm \frac{y - x}{\|y - x\|} \quad \textit{and} \quad \pm \frac{x}{\|x\|}\left(\textit{or} \pm \frac{y}{\|y\|}\right)$$

as the vertices of the unit parallelogram.

The proof is an immediate consequence of the following lemma.

LEMMA. *If ABC is an isosceles triangle in a two dimensional real normed linear space with $\|AB\| = \|AC\|$ and X any point on side AC, then $\|BX\| \geqq \frac{1}{2}\|BC\|$ with equality holding iff $\|AX\| = \frac{1}{2}\|BC\|$ and the unit circle is a parallelogram.*

Proof. Consider points D and E on sides AC and BC respectively such thnt $\|CD\| = \frac{1}{2}\|BC\|$ and the line DE is parallel to line AB. $\|DE\| = \|DC\| = \frac{1}{2}\|BC\|$. If X is on the closed segment DC, the inequality follows from the triangle inequality with equality possible only if $X \equiv D$. If X is on the closed segment AD, consider the point Y on the segment BE such that the line BX is parallel to the line DY. $\|BX\| \geqq \|DY\| \geqq \|DE\| = \frac{1}{2}\|DC\|$, the middle inequality following from the convexity of spheres. Here again equality holds iff $X \equiv D$.

If $\|BD\| = \|DC\| = \frac{1}{2}\|BC\|$ then B and C together with the reflections of these points in D form the vertices of a parallelogram all of whose boundary points are equidistant from D.

Proof of the theorem. The sufficiency is clear. Suppose then that $\|y\| \geqq \|x\| > 0$ and the vectors are linearly independent. The case of linear dependence leads easily to the condition $x = y$. Application of the lemma to the triangle defined by the null vector and the vectors y and $\|y\|x/\|x\|$ implies that

$$\|y - x\| \geqq \tfrac{1}{2}\left\| y - \frac{\|y\|}{\|x\|} x \right\|$$

which is equivalent to the M-S inequality. The lemma further implies that if equality holds then

$$\pm \frac{x}{\|x\|} \quad \text{and} \quad \pm \frac{y-x}{\|y-x\|}$$

are the vertices of the unit parallelogram of the real span of x and y.

A FUNCTIONAL INEQUALITY

SEYMOUR HABER, National Bureau of Standards, Washington

Let f be a continuous and strictly increasing function, defined to the right of some positive number and not bounded from above. Its inverse then has all those properties, and the graph of f^{-1} is the reflection in the line $y=x$ of the graph of f. We will compare the product $f(x)\cdot f^{-1}(x)$ with x^2.

Either may be larger; as for example when f is x^2 and when f is $x-1$. There can of course be no restriction on how much larger than x^2 the product $f(x)\cdot f^{-1}(x)$ might be. The general result in the opposite direction is:

THEOREM 1. *If f is continuous, strictly increasing, and $f(x)\to\infty$ as $x\to\infty$, then for any $\epsilon>0$ and any K*

$$f(x)\cdot f^{-1}(x) > (1-\epsilon)x^2 \qquad \textit{for some } x > K. \tag{1}$$

Proof. We first note that it is sufficient to prove the theorem under the additional assumption that $f(x)\leqq x$ always. For, if f does not satisfy this condition then we may define a function h by $h(x)=f(x)$ when $f(x)\leqq x$, and $h(x)=f^{-1}(x)$ when $f(x)>x$. Since f is continuous, the set of points at which $h(x)=f^{-1}(x)$ is a union of open intervals (a, b) with $f(a)=a$ and $f(b)=b$. On each such interval h is continuous and strictly increasing, and $a<h(x)<b$. It follows that h satisfies all the hypotheses of the theorem; and $h(x)\leqq x$. Furthermore, $h(x)\cdot h^{-1}(x)=f(x)\cdot f^{-1}(x)$, so that f satisfies (1) exactly where h does.

We now assume that $f(x)\leqq x$. Setting $g(x)=f(x)/x$ and $y=f^{-1}(x)$, (1) takes the form

$$g(x)/g(y) \geqq 1-\epsilon. \tag{2}$$

Let $A=\limsup_{x\to\infty} g(x)$. There exists an increasing sequence $\{x_i\}$ going to infinity, such that for each i, $g(x_i)>A-1/i$ and $g(x)<A+1/i$ for all $x\geqq x_i$. Since $y_i=f^{-1}(x_i)\geqq x_i$, we then have

$$g(x_i)/g(y_i) \geqq \frac{A-1/i}{A+1/i}$$

and the theorem follows.

We can say more about the set of points at which (1) holds. Let ϵ be any positive number. Then with x_i as above, for any x in the interval $(x_i, x_i(1+\epsilon/2))$, it is true that

$$g(x) = f(x)/x > f(x_i)/x > \frac{1}{1+\epsilon/2} h(x_i)$$

so that

$$g(x)/g(y) > \frac{1}{1+\epsilon/2} \frac{A - 1/i}{A + 1/i}$$

and so $g(x)/g(y) > 1-\epsilon$ if i is sufficiently large. The lengths of these intervals increase as i increases; so we may conclude:

THEOREM 1′. *If f is as in Theorem* 1, *then for any $\epsilon > 0$ and any K, the set of all $x > K$ for which the inequality* (1) *holds is of infinite* (*Lebesgue*) *measure.*

This seems to be the most that can be said in general; there is a function f such that for any $\delta > 0$, the set of x in $(0, a)$ at which (1) is satisfied is of measure less than δa, for large a. However, imposing a condition of smoothness of growth on f enables us to strengthen the conclusion.

THEOREM 2. *If f satisfies the hypotheses of the above theorem, and $f(x)/x$ is nondecreasing, then for any $\epsilon > 0$ there is a K' such that the inequality* (1) *holds for all $x > K'$; if in addition* $\lim_{x\to\infty} f(x)/x > 1$, *the quantity* $1-\epsilon$ *in* (1) *may be replaced by* 1.

Proof: Define g as above; it is now nondecreasing. Letting $y = f^{-1}(x)$ again, we are to prove that $g(x)/g(y) \geqq 1-\epsilon$. If $g(x) \geqq 1$ from some point on, $f(x) \geqq x$ and so $y \leqq x$. Since g is nondecreasing we would then have $g(x)/g(y) \geqq 1$. If $g(x)$ is always < 1, then $y \geqq x$, so that $g(x)/g(y) \leqq 1$. However if we set $h(x) = f^{-1}(x)/x$, our inequality becomes $h(y)/h(x) \geqq 1-\epsilon$; and h is nonincreasing and bounded below, so that the inequality must hold.

By considering $f^{-1}(x)$ it is seen that (1) also holds for all sufficiently large x if $f(x)/x$ is nonincreasing; and then if $\lim f(x)/x < 1$, (1) holds with $1-\epsilon$ replaced by 1.

A CONDITION FOR A CANCELLATION SEMIGROUP TO BE A GROUP

TREVOR EVANS, Emory University

We describe a rather unusual type of condition for a cancellation semigroup to be a group, in terms of identities which the semigroup may satisfy.

THEOREM. *If a cancellation semigroup satisfies an identity which is not a consequence of the commutative law, then it is a group.*

We actually prove a rather more general result:

If a semigroup satisfies an identity which is not a consequence of the commutative law, then it is periodic (in fact, satisfies an identity $x^m = x^n$, $m \neq n$).

The theorem follows from this and the trivial fact that a periodic cancellation semigroup is a group.

We begin by classifying those semigroup identities which are consequences of the commutative law. If $w(x, y, z, \cdots)$ is a semigroup word in the variables $x, y, z, \cdots$, we will write $\ell(w)$ for the length of w and $\ell_x(w)$ for the number of occurrences of the variable x in w.

LEMMA 1. *In a semigroup, an identity $u(x, y, z, \cdots) = v(x, y, z, \cdots)$ is a consequence of the commutative law if and only if $\ell_x(u) = \ell_x(v)$, for every variable x occurring in the identity.*

Proof. Let $u = v$ be a consequence of the commutative law and let x be any variable occurring in the identity. Since $u = v$ holds in every commutative semigroup, it holds in the cyclic group G of prime order p (considered as a semigroup). Put all variables in u, v other than x, equal to the neutral element 1 in G. Then

$$x^{\ell_x(u)} = x^{\ell_x(v)}, \quad \text{for all } x \text{ in } G.$$

Hence, $\ell_x(u) \equiv \ell_x(v) \pmod{p}$.

Since this congruence holds for every prime, $\ell_x(u) = \ell_x(v)$. Conversely, if $\ell_x(u) = \ell_x(v)$ for every variable, then the generalized commutative-associative law implies $u = v$. This completes the proof.

LEMMA 2. *Let S be a semigroup satisfying an identity which is not a consequence of the commutative law. Then S is periodic, in fact, satisfies an identity $x^m = x^n$, $m \neq n$.*

Proof. Let S satisfy the identity $u(x, y, z, \cdots) = v(x, y, z, \cdots)$ which is not a consequence of the commutative law. If $\ell(u) \neq \ell(v)$, then putting all variables in $u = v$ equal to x, we see that x satisfies $x^{\ell(u)} = x^{\ell(v)}$ for all x. Hence, S is periodic.

If $\ell(u) = \ell(v)$, a slight modification is needed. By Lemma 1, there is a variable x occurring in the identity such that $\ell_x(u) \neq \ell_x(v)$. Put all variables in $u = v$, other than x, equal to x^2. Then S satisfies the identity

$$x^{2\ell(u) - \ell_x(u)} = x^{2\ell(v) - \ell_x(v)}, \quad \text{for all } x.$$

Again, S is periodic.

LEMMA 3. *A periodic cancellation semigroup is a group.*

This well-known result [1, p. 23, Ex. 6(c)] is an immediate consequence of the fact that a local one-sided neutral element in a cancellation semigroup is a two-sided neutral element for the whole semigroup.

Since it is known what type of identities a group can satisfy [2], the following form of the theorem

If a cancellation semigroup is not a group, then any identity it satisfies is a consequence of the commutative law

determines the identities a cancellation semigroup can satisfy. More generally, any identity which a nonperiodic semigroup satisfies is a consequence of the commutative law. In this form we see that our result is, in a way, a generalization of Neumann's result in [2], that any identity which a group satisfies may be put in the form

$$x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} \cdot c(x_1, x_2, \cdots, x_n) = 1,$$

where $c(x_1, x_2, \cdots, x_n)$ is a complex commutator word in the variables.

References

1. A. H. Clifford and G. B. Preston, The algebraic theory of semigroups, vol. 1, Math. Surveys 7, Amer. Math. Soc., 16 (1962).
2. B. H. Neumann, Identical relations in groups I, Math. Ann., 114 (1937) 506–525.

A NOTE ON FERMAT'S LAST THEOREM

J. M. Gandhi, University of Alberta, Edmonton

Recently D. E. Stone [2] proved the following theorem which pertains to Fermat's Last Theorem (FLT).

THEOREM. *If p and $2p+1$ are odd primes and*

$$a^p + b^p + c^p = 0,$$

where a, b, c are nonzero, pairwise prime integers, then precisely one of the integers a, b, c is divisible by p.

In this note we prove a similar theorem pertaining to FLT.

THEOREM. *If p and $4p+1$ are primes with $p>3$, and $a^p+b^p+c^p=0$, where a, b, c are nonzero, pairwise prime integers, then precisely one of the integers a, b, c is divisible by $4p+1$.*

Proof of Theorem. Assume that $(abc, 4p+1)=1$. Writing (1) as $a^p+b^p=-c^p$ and squaring, we get

$$a^{2p} + b^{2p} + 2a^pb^p = c^{2p}. \tag{1}$$

Since $4p+1$ is prime, and since, by assumption, a, b, c are each prime to $4p+1$, we have by Fermat's "little" theorem,

$$\begin{aligned} &a^{2p} \equiv \pm 1 \pmod{4p+1}, \quad b^{2p} \equiv \pm 1 \pmod{4p+1} \\ &c^{2p} \equiv \pm 1 \pmod{4p+1}. \end{aligned} \tag{2}$$

The sign before the residues is $+$ or $-$ according as $k^2 \equiv a$, $k^2 \equiv b$, $k^2 \equiv c$ (mod $4p+1$) respectively has or has not integral solutions.

Using (2) in (1) we get

$$\pm 1 \pm 1 + 2a^p b^p \equiv \pm 1 \pmod{4p+1}$$

or

$$2a^p b^p \equiv +1, -1, +3 \text{ or } -3 \pmod{4p+1}.$$

Squaring the last congruence and using (2) we get

$$\pm 4 \equiv 1 \text{ or } 9 \pmod{4p+1}.$$

Since p is a prime >3, the last congruence is impossible. This contradicts the assumption $(abc, 4p+1)=1$ and hence one of a, b, c must be divisible by $4p+1$.

The author is on leave from University of Rajasthan, Jaipur.

References

1. L. E. Dickson, History of Theory of Numbers, Vol. II, Chelsea, New York, 1952, 734.
2. D. E. Stone, On Fermat's Last Theorem, this MONTHLY, 70 (1963) 976.

CLASSROOM NOTES

EDITED BY GERTRUDE EHRLICH, University of Maryland

Send manuscripts to R. A. Rosenbaum, Wesleyan University, Middletown, Conn. 06457.

ON PROVING THEOREMS IN PLANE GEOMETRY VIA DIGITAL COMPUTER

RICHARD BELLMAN, University of Southern California

1. Introduction. The development of the digital computer has focused considerable attention upon various types of algorithms, and, in particular, upon those connected with logical processes and decision-making. An offshoot of this has been the set of attempts by various people, with varying degrees of success, to replicate human thought processes with the aid of a digital computer. In this connection, let us cite the work in translation of languages, pattern recognition, chess playing, checker playing, and the proving of logical and geometric theorems.

In pursuing these goals, there are many different approaches that can be pursued. At one extreme, we can imitate what the human mind does; at the other extreme, we can fasten our attention solely upon the capabilities of an analog or digital computer. In between, we have a continuum of man-machine processes. Since it is generally agreed by knowledgeable people that we possess very little understanding of the working of the brain, it is clearly hazardous to follow the first route. We shall restrain ourselves exclusively to a completely rigorous use of the computer in establishing geometric theorems.

The basic idea is quite simple. The structure of Euclidean plane geometry permits us to express geometric theorems as algebraic identities. Algebraic iden-

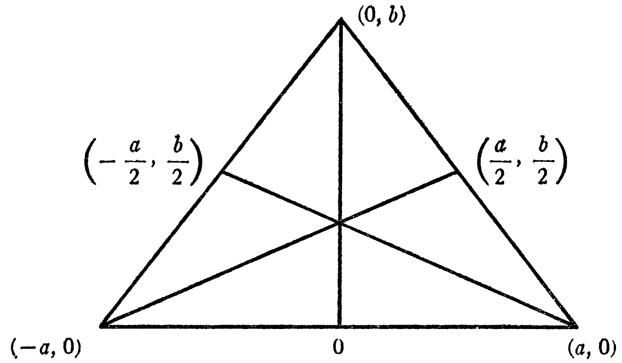
tities can be established by verification of a sufficiently large number of cases. This verification should be possible, purely arithmetically, using a digital computer. There are, however, many interesting problems associated with a procedure of this type, as we shall see below.

Let us hasten to add that we see no intrinsic value to establishing geometric theorems in this fashion. We do feel that it is pedagogically of some value to have the student interested in the uses of the computer try his skill at problems of this type, and there are some nontrivial associated arithmetic and analytic questions. Furthermore, numerous questions arise as to what we mean by "fundamental theorems" of geometry. All of this will be illustrated by the following discussion.

2. The medians of an isosceles triangle. Suppose that we wish to prove that the medians to the equal sides of an isosceles triangle are equal. Considering the figure below, we must establish the identity

$$\begin{aligned}(2.1)\qquad &\left[\left(a-\left(-\frac{a}{2}\right)\right)^2+\left(0-\frac{b}{2}\right)^2\right]^{1/2}\\ &\qquad=\left[\left(\frac{a}{2}-(-a)\right)^2+\left(\frac{b}{2}-0\right)^2\right]^{1/2}.\end{aligned}$$

This is, of course, obvious upon inspection, but with a digital computer, we are not allowed "inspection."



As far as the computer is concerned, an isosceles triangle is determined by the three vertices $(0, b)$, $(a, 0)$, $(-a, 0)$, and the lengths of the medians are determined by simple algorithms which provide first the midpoints and then the length according to the distance formula. We are assuming the Pythagorean theorem which enables us to use the usual distance formula of analytic geometry.

We establish the equality of the two sides of (2.1) in a two-step process. We first invoke some general algebraic theorems which assure us that it is sufficient to verify equality for a finite set of values of a and b, and then use the computer to carry out this arithmetic confrontation.

There are several ways to proceed. Let us sketch one. Since the expressions are homogeneous in a and b (scale is unimportant), it suffices to take $b=1$. Since equality of the squares implies equality of positive quantities, let us square both sides. It remains to establish the identity of two quadratic polynomials in a. For this, equality at *three* values of a suffices. Choose three convenient values, e.g., $a=2, 4, 6$.

At this point, the reader may justifiably worry about round-off error. After all, computer arithmetic is not ordinary arithmetic. Suppose that we had not thought of the artifice of squaring, or, in general, of rationalizing. How would we establish that $M=N$ by comparing the calculated values of $\sqrt{M}$ and $\sqrt{N}$? Does agreement of $\sqrt{M}$ and $\sqrt{N}$ to a sufficiently large number of decimal places assure us that they are equal? The answer is "yes."

Observe that

$$\sqrt{M}-\sqrt{N}=\frac{M-N}{\sqrt{M}+\sqrt{N}}\cdot \tag{2.2}$$

Hence, if M and N are *integers* and distinct, we must have

$$\left|\sqrt{M}-\sqrt{N}\right| \geqq \frac{1}{\sqrt{M}+\sqrt{N}}\cdot \tag{2.3}$$

If arithmetic calculations show that

$$\left|\sqrt{M}-\sqrt{N}\right| \leqq \epsilon < \frac{1}{\sqrt{M}+\sqrt{N}}, \tag{2.4}$$

we can conclude that $M=N$. Starting with the values of M and N, we know how to obtain the accuracy of (2.4).

In the general case, a number of interesting questions arise as to the number of verifications required and the methods to be used to obtain this verification. How, for example, does one systematically reduce a problem involving distances to a polynomial identity?

3. Discussion. We leave it to the reader to investigate the possibility of establishing the existence of the Euler line, the Simpson line, the nine-point circle, and so on. It is also clear that we can "generate" theorems of this type in a completely uninspired fashion by tabulating sets of points and lines and testing collinearity, coincidence, etc. Some of this would be pursued in an adaptive fashion, as, for example, the search for the nine-point circle. As mentioned previously, none of this has any intrinsic interest. But these are useful problems for training purposes, since they deal with familiar types of questions requiring no advanced training. They illustrate what is meant by the term "algorithm," and they also demonstrate the value of ingenuity and knowledge of the structure of a process.

FACTORIZATIONS WITH UNEQUAL NUMBERS OF PRIMES

BERNARD JACOBSON, Franklin and Marshall College

Hilbert was probably the first to give a very simple example of a semigroup which failed to possess unique factorization into "primes." In the set $H=\{x: x\equiv 1 \bmod 4, x>0\}$, $441=21\cdot 21=9\cdot 49$ and 9, 21, 49 are "primes" in H. Other examples are given in [3], and [4]. In these examples and in all of the examples in elementary textbooks which have come to the attention of the author, the same number of "primes" appear in each of the different factorizations. On a somewhat higher level it has been shown that the algebraic number field Z has a class number $\leqq 2$ if and only if for every nonzero integer $\alpha\in Z$ the number of primes in every factorization of α into a product of primes depends only on α (see [1]).

A class of examples of different factorizations with unequal numbers of "primes" is set forth below. Here, the product of two "primes" is shown to be equal to the product of n "primes" for any $n>1$. The curious and interested student may wish to try to generalize and show the product of m "primes" equal to the product of n "primes."

An element z is called prime in the set $S=\{x: x\equiv 1 \bmod p, x>0\}$ if and only if $z=uv$ with $u, v\in S$ implies $z=u$ or $z=v$.

THEOREM. *Let n be a positive integer, p a rational prime such that $p\equiv 1 \bmod n$, and $S=\{x: x\equiv 1 \bmod p, x>0\}$, then there exists an integer k in S such that $k=rs=t^n$ where r, s, t are prime in S.*

Proof. From $p\equiv 1 \bmod n$, $n|p-1$ $(=\phi(p))$, and thus there exist $\phi(n)$ residue classes of order n modulo p. Let a belong to n, i.e. $a^n\equiv 1 \bmod p$ and $a^m\equiv 1 \bmod p$ implies $n|m$. Then $a^i\not\equiv 1 \bmod p$ for $i=1, 2, \cdots, n-1$. Since $(a, p)=1$, there exists b such that $ab\equiv 1 \bmod p$. Clearly $b^n\equiv 1 \bmod p$ and $b^i\not\equiv 1 \bmod p$ for $i=1, 2, \cdots, n-1$. By Dirichlet's theorem there exist rational primes q_1 and q_2 such that $q_1\equiv a \bmod p$ and $q_2\equiv b \bmod p$. It now follows that $(q_1q_2)^n=(q_1^n)(q_2^n)$, where q_1^n, q_2^n and q_1q_2 are primes in S.

Examples. (1) Let $n=3$, $p=7$, $S=\{1, 8, 15, \cdots\}$, $q_1=2$ and $q_2=11$, then $10648=22^3=8\cdot 1331$.

(2) Let $n=4$, $p=13$, $S=\{1, 14, 27, \cdots\}$, $q_1=5$, $q_2=47$, then $235^4=(5^4)(47^4)$.

(3) Let $n=5$, $p=11$, $S=\{1, 12, 23, \cdots\}$, $q_1=3$, $q_2=37$, then $111^5=(3^5)(37^5)$.

References

1. L. Carlitz, A characterization of algebraic number fields with class number two, Proc. Amer. Math. Soc., 11 (1960) 391–392.

2. Harvey Cohn, A Second Course in Number Theory, Wiley, New York, 1962.

3. Newcomb Greenleaf and Robert J. Wisner, The unique factorization theorem, Math. Teacher, 52 (1959) 600–603.

4. Bernard Jacobson and Robert J. Wisner, Matrix number theory: an example of nonunique factorization, this MONTHLY, 72 (1965) 399–402.

A REMARK ON CUBICS

L. CARLITZ, Duke University

It is familiar that the discriminant of the cubic $f(x)=x^3+px+q$ is given by $D=-4p^3-27q^2$. In particular the discriminant of x^3-m is equal to $-27m^2$; hence for rational m the discriminant is of the form $-3k^2$, where k is rational.

Now let θ satisfy the cubic equation

$$\theta^3 + p\theta + q = 0, \tag{1}$$

where p and q are rational and put

$$\phi = a + b\theta + c\theta^2. \tag{2}$$

We shall show that, provided

$$4p^3 + 27q^2 = 3k^2, \tag{3}$$

where k is rational, we can find rational a, b, c, not all zero, such that ϕ^3 is rational.

In the first place it follows from (1) that

$$\sum \theta = 0, \qquad \sum \theta^2 = -2p, \qquad \theta\theta'\theta'' = -q, \tag{4}$$

where θ, θ', θ'' are the roots of (1). Then by (2) and (4)

$$\sum \phi = 3a - 2cp. \tag{5}$$

In the next place we have

$$\phi^2 = (a^2 - 2bcq) + (2ab - 2bcp - c^2q)\theta + (b^2 + 2ac - c^2p)\theta^2,$$

so that

$$\sum \phi^2 = 3(a^2 - 2bcq) - 2p(b^2 + 2ac - c^2p). \tag{6}$$

We now require that

$$\sum \phi = \sum \phi^2 = 0. \tag{7}$$

In view of (5) and (6), (7) implies

$$c^2p^2 - 9bcq - 3b^2p = 0. \tag{8}$$

We may think of (8) as a quadratic in b, c; the discriminant is equal to

$$12p^3 + 81q^2 = 3(4p^3 + 27q^2) = 9k^2,$$

by (3). The roots of (8) are therefore rational. We find that

$$\frac{c}{b} = \frac{3(3q \pm k)}{2p^2}.$$

There is no loss in generality in taking the upper sign. Then by (5)

$$a = \frac{3q + k}{p} b, \qquad c = \frac{(3(q + k)}{2p^2} b. \tag{9}$$

To compute ϕ^3, we put $\phi^2 = A + B\theta + C\theta^2$, where

$$\begin{cases} A = a^2 - 2bcq, \\ B = 2ab - 2bcp - c^2q, \\ C = b^2 + 2ac - c^2p. \end{cases}$$

Then $\phi^3 = (aA - bqC - cqB) + (aB + bA - bpC - cpB - cqC)\theta + (aC + bB + cA - cpC)\theta^2$. In view of (7) this reduces to

$$\phi^3 = aA - bqC - cqB. \tag{10}$$

Making use of (9) we find after a little manipulation that (10) becomes

$$\phi^3 = \frac{6(3q + k)^2 k^3 b^3}{8p^6}. \tag{11}$$

Conversely if we put

$$\begin{cases} \theta = a \cdot 6^{1/3}(3q + k)^{2/3} + b \cdot 6^{2/3}(3q + k)^{1/3}, \\ \theta' = a\omega \cdot 6^{1/3}(3q + k)^{2/3} + b\omega^2 \cdot 6^{2/3}(3q + k)^{1/3}, \\ \theta'' = a\omega^2 \cdot 6^{1/3}(3q + k)^{2/3} + b\omega \cdot 6^{2/3}(3q + k)^{1/3}, \end{cases} \tag{12}$$

where ω, ω^2 are the primitive cube roots of unity, we find that

$$\sum \theta = 0, \quad \sum \theta^2 = 36ab(3q + k), \quad \theta\theta'\theta'' = 6(3q + k)^2 a^3 + 36(3q + k)b^3.$$

We now require that

$$\sum \theta^2 = -2p, \qquad \theta\theta'\theta'' = -q. \tag{13}$$

We find without much difficulty that (13) is satisfied provided

$$a = \frac{p}{3(3q + k)}, \qquad b = -\frac{1}{6}. \tag{14}$$

An algebraic number field of the type $R(\sqrt[3]{m})$, where m is rational but not the cube of a rational number, may be called a pure cubic field. The discriminant of such a field is evidently of the form $-3k^2$, where k is some rational integer. As an application of the above discussion we may state the following

THEOREM. *A cubic field is pure if and only if its discriminant is of the form* $-3k^2$, *where* k *is some positive rational integer.*

A THEOREM ON GROUPS AND THE CHARACTERISTIC OF AN INTEGRAL DOMAIN

R. W. BALL, Auburn University

If the characteristic of an integral domain is finite, it is prime. The standard textbook proofs, using the absence of zero-divisors, seem to miss the point in that they cannot be generalized, for example, to simple rings as suggested by Jacobson, Lectures in Abstract Algebra, volume 1. The purpose of this note is to call attention to the group-theoretic nature of this result.

Let G be a group with identity e.

THEOREM. *If $G \neq \{e\}$ and all elements of G other than e have the same finite order m, then m is a prime.*

Proof. Assume that m is composite, say $m=rs$ with $1<r<m$. Let a be an element of G, $a\neq e$. Then $(a^s)^r=a^m=e$. Thus $a^s=e$ since $r<m$, and $a=e$ since $s<m$. This is a contradiction.

In the additive notation this equation is $r(sa)=ma=0$.

THEOREM. *If R is a ring with at least two elements in which all nonzero elements have the same additive order m (so that m is the characteristic of R), then either m is infinite or it is a prime.*

Using the absence of zero-divisors, it can be proved that an integral domain satisfies the hypothesis of this theorem. Similarly a simple ring (ring with no nontrivial ideals) has this property since for each positive integer k, $\{a \text{ in } R \mid ka=0\}$ is a (two-sided) ideal of R. In general if R contains an element $a\neq 0$ of minimal additive order such that either $R=aR$ or $R=Ra$, then R is homogeneous in the sense of the above theorem. An identity 1 in an integral domain is such an element.

A CHARACTERIZATION OF FINITE NILPOTENT GROUPS

C. V. HOLMES, San Diego State College

It will be shown that a finite group is nilpotent if and only if there is at least one normal subgroup of order m for each positive integer m which is a divisor of the order of the group. A finite Hamiltonian group is nilpotent and this characterization of finite nilpotent groups classes them as being near Hamiltonian.

A group G is said to be nilpotent if there exists a positive integer n such that $Z_n=G$, where Z_n is a member of the upper central series, $E=Z_0, Z_1, Z_2, \cdots$ of G, where

$$Z_{n+1} = \{x \in G \mid [x, y] \in Z_n \quad \text{for all } y \in G\}.$$

For finite groups the following are equivalent:

(1) G is nilpotent.

(2) If H be a proper subgroup of G then H is a proper subgroup of the normalizer of H.

(3) G is a direct sum of its Sylow subgroups.

We extend this list with the following theorem.

THEOREM 1. *A finite group G is nilpotent if and only if for each divisor d of the order of G there exists a normal subgroup of order d.*

Proof. Let G be a group of order $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$, p_i a prime, and let d be a divisor of n, $d=p_1^{\beta_1}p_2^{\beta_2}\cdots p_k^{\beta_k}$, $0\leqq\beta_i\leqq\alpha_i$.

Suppose that G is a nilpotent group. Write G as a direct sum of its Sylow-p subgroups, $G=\sum_{i=1}^k(S(p_i))$, $S(p_i)$ the Sylow-p_i subgroup of G. $S(p_i)$ is a p-group of order $p_i^{\alpha_i}$, and hence $S(p_i)$ contains a normal subgroup H_i of order $p_i^{\beta_i}$. Since $S(p_i)$ is both normal in G and, moreover, a summand of G it follows that H_i is normal in G. That H_i is disjoint from the group generated by $H_1, \cdots, H_{i-1}, H_{i-1}, \cdots, H_k$ follows since H_i is a subgroup of $S(p_i)$ and the subgroups $S(p_j)$ enjoy this disjointness property. Then $H=\sum_{i=1}^k H_i$ exists as a subgroup of G, and has order d. That H is a normal subgroup of G follows from the fact that each H_i is normal in $S(p_i)$.

Conversely, since $p_i^{\alpha_i}$ is a divisor of n, G has a normal subgroup of order $p_i^{\alpha_i}$ which is the unique Sylow-p_i subgroup of G. Hence $G=\sum_{i=1}^k S(p_i)$, and we conclude that G is nilpotent.

A PROOF OF THE EQUALITY OF COLUMN AND ROW RANK OF A MATRIX

HANS LIEBECK, University of Keele, Staffordshire, England

Let A be an $m\times n$ complex matrix, A^* its conjugate transpose, and let x and y denote $n\times 1$ matrices (column vectors).

LEMMA 1. *$y^*y=0$ if and only if $y=0$.*

LEMMA 2. *$Ax=0$ if and only if $A^*Ax=0$.*

Proof. If $A^*Ax=0$, then $y^*y=0$, where $y=Ax$. Hence, by Lemma 1, $y=0$. The converse is obvious.

Now let $R(A)$ denote the range of A, i.e. the vector space $\{Ax;\text{ all } x\}$. Note that column rank (c.r.) $A=\dim R(A)$. We write r.r. A for the row rank of A.

LEMMA 3. $\dim R(A)=\dim R(A^*A)$.

Proof. By Lemma 2, $Ax_1, \cdots, Ax_k$ are linearly independent if and only if $A^*Ax_1, \cdots, A^*Ax_k$ are linearly independent.

THEOREM. c.r. $A=$c.r. $A^*=$r.r. A.

Proof. c.r. $A=$c.r. $A^*A=\dim\{A^*(Ax);\text{ all } x\}\leqq\dim\{A^*y;\text{ all } y\}=$c.r. A^*. Thus also c.r. $A^*\leqq$c.r. $A^{**}=$c.r. A, and so c.r. $A=$c.r. $A^*=$r.r. $\overline{A}=$r.r. A.

PROBLEMS AND SOLUTIONS

EDITED BY E. P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, N. J., 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of ElementaryProblems should be sent to M. S. Klamkin, Ford Scientific Laboratory, P. O. Box 2053, Dearborn, Mich. 48121. To facilitate their consideration, solutions for Elementary Problems in this issue should be submitted on separate, signed sheets and should be mailed before April 30, 1967.

E 1935. *Proposed by W. J. Blundon, Memorial University of Newfoundland*

Prove that, for every triangle ABC, $s \leqq 2R+(3\sqrt{3}-4)r$, where s, R and r are the semiperimeter, circumradius and inradius, respectively. Equality holds only for the equilateral triangle.

E 1936. *Proposed by W. J. Blundon, Memorial University of Newfoundland*

If in triangle ABC we have

$$\frac{\sin A + \sin B + \sin C}{\cos A + \cos B + \cos C} = \sqrt{3},$$

prove that at least one angle of the triangle is 60°.

E 1937. *Proposed by J. M. Quoniam, Saint-Etienne, France*

Taking

$$S_\mu = \sum_{k=1}^{[n/2]} 4^\mu \cos^{2\mu} \frac{k\pi}{n+1},$$

prove that

$$S_\mu = (n+1)\binom{2\mu-1}{\mu-1} - 2^{2\mu-1},$$

where the usual symbols for greatest integer and binomial coefficient have been used.

E 1938. *Proposed by Stanimir Fempl, University of Belgrade, Yugoslavia*

A right cone and a right circular cylinder have a common base and altitude. If a plane meets these two bodies so that the intersections are ellipses, prove that the ratio of the eccentricities of the ellipses is equal to the ratio of the lengths of the generatrices of the bodies.

E 1939. *Proposed by Richard Stanley, California Institute of Technology*

How many of the first n triangular numbers $T_k = \frac{1}{2}k(k+1)$ are divisible by n?

E 1940. *Proposed by J. V. Cornacchio and R. P. Soni, IBM, Endicott, N. Y.*

Given the finite sequence $\{p_k\}_{k=1}^{m}$ such that $0 < p_k \leqq 1$ $(k=1, 2, \cdots, m)$, $\sum_{k=1}^{m} p_k = 1$, and the $m \times m$ matrix $\|\alpha_{kl}\|$, each of whose elements satisfies the condition $|\alpha_{kl}| \leqq 1$, prove (or disprove):

$$\sum_{k=1}^{m}\sum_{l=1}^{m} p_k p_l |\alpha_{kl}|^2 = 1 \text{ if and only if } |\alpha_{kl}| = 1, \qquad (k, l = 1, 2, \cdots, m).$$

E 1941. *Proposed by William Koenen, Highland Park High School, St. Paul, Minn.*

Let M be the set of natural numbers. For $A \subseteq M$, we let cA represent the complement of A in M.

(1) If kA is the smallest set containing A which is closed for addition, prove that there are no more than six sets of the form A, cA, kA, ckA, kcA, $ckcA$, $kckA$, $ckckA$, $kckcA$, etc., and display a set A for which six are distinct.

(2) If hA is the smallest set containing A which is closed for multiplication, display a set A for which fourteen of the sets, A, cA, hA, chA, hcA, $chcA$, $hchA$, $chchA$, $hchcA$, $chchcA$, etc., are distinct.

E 1942. *Proposed by C. F. McLaren, University of Michigan*

Let

$$A = \begin{pmatrix} -91 & 28 & 9 \\ 47 & -14 & -4 \\ -1113 & 341 & 108 \end{pmatrix} \cdot \quad \text{Find } \sqrt{A}.$$

E 1943. *Proposed by J. M. Khatri, Baroda, India*

(1) Prove or disprove: There exists an infinite series of triangular numbers such that every partial sum is a perfect square number. (2) The same except that every partial sum shall be a triangular number.

E 1944. *Proposed by Philip Dwinger, University of Illinois at Chicago*

Let X be a set of n points, $n \geqq 1$. Let $A_1, A_2, \cdots, A_k$ be a family of k subsets of X, such that every subset of X can be expressed in terms of the A_i, $1 \leqq i \leqq k$, by means of the set-theoretic operations: union, intersection and complementation. Find the minimum value of k.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Rutgers—The State University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be submitted on separate, signed sheets and should be mailed before June 30, 1967.

5440. *Proposed by Benjamin Volk, Yeshiva University*

Let $F(z)$ be a complex-valued function of a complex variable defined and bounded for $|z|<1$. Prove that the following are equivalent:

(A) $F(z)$ is analytic for $|z|<1$ and absolutely monotonic in $(-1, 1)$.

(B) For any sequence $\{z_n\}$ of distinct points in $\{z: |z|<1\}$ we have

$$\left| [z_1, z_2, \cdots, z_k] \right| \leqq \left| [|z_1|, |z_2|, \cdots, |z_k|] \right|, \quad k = 1, 2, \cdots,$$

and $[|z_1|, |z_2|, \cdots, |z_k|]$ is monotonic in $|z_1|$. Here we use the notation for a given function $F(x)$ as defined in Milne-Thompson, *The Calculus of Finite Differences*, p. 1.

5441. *Proposed by Benjamin Volk, Yeshiva University*

Find a real-valued function $f(x)$ of a real variable such that

(A) $f(x)$ is Lebesgue integrable on $[\delta, 1]$ for every $\delta>0$,

(B) $\operatorname{Lim}_{x\to\infty} (1/x) \sum_{k\leqq x} f(k/x)$ exists,

(C) $\operatorname{Lim}_{\delta\to 0} \int_\delta^1 f(t)dt$ does not exist.

5442. *Proposed by J. T. Renfrow, California Institute of Technology*

Let $f(x)$ be an infinitely differentiable real-valued function defined on an open interval (a, b). If $f(x)$ vanishes infinitely often on a closed bounded subinterval $[c, d]$, $c\neq d$, and $\sup_{x\in(a,b)}|f^{(n)}(x)| = O(n!)$ as $n\to\infty$, then $f(x)$ vanishes identically on an open subinterval of (a, b).

5443. *Proposed by G. E. Noether, Boston University*

Let $x_{(1)}, \cdots, x_{(n)}$ be an ordered sample from a normal distribution, and let $\xi_{i/n+1}$ be the ith $(n+1)$-ile of this distribution. Prove that for $i>\frac{1}{2}n$, $\xi_{i/n+1} \leqq E(x_{(i)}) \leqq \xi_{i+1/n+1}$.

5444. *Proposed by T. S. Frank, Le Moyne College, and W. A. Lopez, University of Wisconsin*

It is well known that the intersection of compact sets in a Hausdorff space is again compact. Prove or disprove the corresponding statement for T_1 spaces.

5445. *Proposed by A. J. Macintyre and C. I. Lubin, University of Cincinnati*

Show that

$$e = 1 + \frac{2}{1+}\,\frac{1}{6+}\,\frac{1}{10+}\,\frac{1}{14+}\,\frac{1}{18+}\cdots$$

and the sequence of convergents to this fraction form a subsequence of the se-

quence of convergents of the regular fraction for e. Show also that

$$e^2 = 7 + \frac{2}{5+}\,\frac{1}{7+}\,\frac{1}{9+}\,\frac{1}{11+}\cdots$$

and the sequence of convergents of this fraction form a subsequence of the sequence of convergents of the regular fraction for e^2.

5446. *Proposed by R. Sivaramakrishnan, Engineering College, Trichur, India*

Let $\lambda(n)$ denote the number of distinct prime divisors of n and let $\Omega(n)$ be the total number of prime divisors of n. Prove that

$$\sum_{d|n} (-1)^{\Omega(d)}\phi(d)\sigma(n/d) = \sum_{k^2|n} k^2\cdot 2^{\lambda(n/k^2)},$$

where $\phi(n)$ is the Euler phi-function and $\sigma(n)$ is the sum of the divisors of n.

5447. *Proposed by H. A. Smith, Institute for Defense Analysis*

Let K be a compact subset of the plane having a nonvacuous interior. Let A be a linear transformation on the plane which maps K onto itself. Show that if A does not preserve orientation, then A^2 is the identity.

5448. *Proposed by H. A. Smith, Institute for Defense Analysis*

Let A be an $n\times n$ matrix, $A^2=I$, the identity, and $A\neq \pm I$. Show

(1) $\quad \mathrm{Tr}(A) \equiv n \pmod 2$, $\qquad$ (2) $\quad |\mathrm{Tr}(A)| \leqq n-2$,

where $\mathrm{Tr}(A)$ is the trace of A.

5449. *Proposed by G. R. Sell, University of Minnesota*

Let $f: K\times R\to R$ be a bounded, continuous function, where R denotes the real numbers and K is a compact set in R. It can be easily shown that if f is uniformly continuous, then f satisfies the following condition:

(U): For every $\epsilon>0$ there is a $\delta>0$ such that

$$|f(x,t) - f(x,\tau+t)| \leqq \epsilon, \qquad ((x,t)\in K\times R),$$

whenever $|\tau|\leqq\delta$. Is the converse true? That is, does condition (U) imply that f is uniformly continuous?

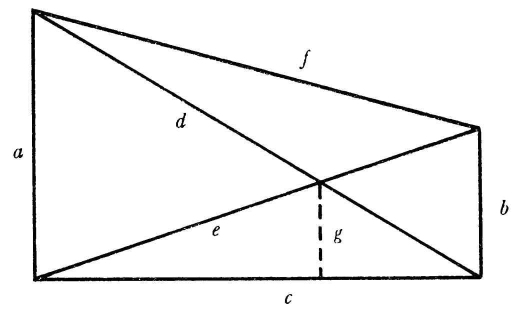
SOLUTIONS OF ADVANCED PROBLEMS

Integral Solutions of the Ladder Problem

5323 [1965, 914]. *Proposed by Alan Sutcliffe, Congleton, Cheshire, England*

There is a well-known problem in which two ladders rest between two walls. The lengths of the ladders and the height above the ground at which they cross are given, and it is required to find the distance between the two walls and the heights that the ladders reach up the walls. The complete solution when all these

six quantities are integers is known (Mathematical Gazette, 47 (1963) 133–136). Is there a solution in which the distance between the tops of the ladders is also integral, other than the trivial case in which the ladders are of equal length?



Solution by Gerald J. Janusz, Institute for Advanced Study and University of Chicago. If, as shown in the figure, a, b, c, d, e, f are integers, then $g=ab/(a+b)$ is rational and a solution is obtained. Thus we seek integral solutions to the system:

$$a^2+c^2=d^2, \qquad b^2+c^2=e^2, \qquad (a-b)^2+c^2=f^2. \tag{1}$$

We can obtain a solution to (1) if there are integers u, v, r, s, z, w such that

$$uv=rs=zw=(c/2)^2, \qquad (u-v)-(r-s)=(w-z), \tag{2}$$

for we then take $a=u-v$, $b=r-s$, whence $a-b=w-z$, $d=u+v$, $e=r+s$, $f=z+w$.

Suppose that m, n, p, q are integers such that

$$q^2=mn(mn+p^2)[(m+n)^2+p^2]. \tag{3}$$

Setting $u=(mn+p^2)(m+n)^2m$, $s=(mn+p^2)(m+n)^2n$,

$$z=(mn+p^2)(m+n)p^2, \qquad uv=rs=zw=p^2(m+n)^2q^2,$$

we obtain a solution for (2).

There are infinitely many solutions to (3). For example, take x any positive integer and set

$$p=2x+1 \quad m=x(2x+1), \quad n=x, \; q=x(x+1)(2x+1)(2x^2+2x+1).$$

After removing a common factor $x(2x+1)$, this gives:

$$\begin{aligned}
a &= (2x+1)(2x^3+6x^2+4x+1)(2x^3+2x^2-1),\\
b &= (6x^3+10x^2+6x+1)(2x^3+2x^2+2x+1),\\
c &= 4x(x+1)^2(2x+1)(2x^2+2x+1),\\
d &= (2x+1)[4x^2(x+1)^4+(2x^2+2x+1)^2],\\
e &= (2x+1)^2(2x^2+2x+1)^2+4x^2(x+1)^4\\
f &= 2(x+1)[x^2(2x^2+2x+1)^2+(x+1)^2(2x+1)^2].
\end{aligned}$$

There is no common factor (>1) for this set for any value of x. In fact for each x, b and c are relatively prime.

If we set $x=1$ and multiply all terms by $2\cdot 139$ we get the following solution to the original problem:

$$a = 2\cdot 3^2\cdot 13\cdot 139, \qquad b = 2\cdot 7\cdot 23\cdot 139, \qquad c = 2^5\cdot 3\cdot 5\cdot 139,$$

$$d = 2\cdot 3\cdot 89\cdot 139, \qquad e = 2\cdot 17^2\cdot 139, \qquad f = 2^3\cdot 61\cdot 139, \qquad g = 3^2\cdot 13\cdot 7\cdot 23.$$

This solution has the additional property that $(a-b)^2+a^2=(2\cdot 5^3\cdot 139)^2$.

Constructing Tangents to Level Lines of u^pv^q

5340 [1965, 1134]. *Proposed by Necdet Ücoluk, Purdue University*

Let A and B be two fixed points in the plane. If N is a variable point of the plane, say $u=\overline{NA}$, $v=\overline{NB}$, determine a geometric construction for the tangent line at the general point N on the locus of $uv^2=k$.

Solution by R. C. Lyness, Blackpool, England. More generally, $u^pv^q=$ constant if and only if $(p/u)(du/ds)=-(q/v)(dv/ds)$ if and only if $pv\cdot\cos\sphericalangle ANC =qu\cdot\cos\sphericalangle BND$, where CND is the tangent. Construct E to divide AB in the ratio $p:q$. Then $u\cdot\sin\sphericalangle ANE/v\cdot\sin\sphericalangle ENB=p/q$, and the isogonal conjugate of EN with respect to AN, BN is the normal at N.

Also solved by J. M. Quoniam (France), and G. D. Kyle.

Sum-Integral Analogs for sinh x

5341 [1965, 1134]. *Proposed by George Purdy, The University, Reading, England*

It is known that

$$\frac{1}{2}\sum_{t=1}^{2}(-1)^te^{(-1)^tx}=\sinh x.$$

Prove the integral analog:

$$\frac{1}{2}\int_1^2(-1)^te^{(-1)^tx}dt=\frac{\sinh x}{i\pi x}\cdot$$

I. *Solution by Michael Schulz, Massachusetts Institute of Technology.* Replace $(-1)^t$ by $e^{(2n+1)i\pi t}$, where n is an integer. Then

$$\frac{1}{2}\int_1^2 e^{(2n+1)i\pi t}\exp(e^{(2n+1)i\pi t}x)dt=\frac{1}{2x(2n+1)i\pi}\left[e^x-e^{-x}\right]=\frac{\sinh x}{(2n+1)i\pi x}\cdot$$

If only the $n=0$ branch is considered, we obtain the desired result; the case $x=0$ may be verified directly by using 1 for sinh 0/0.

II. *Solution by Harsh Pittie, Swarthmore College.* Let $u=(-1)^t=e^{i\pi t}$. Then,

$$\frac{1}{2}\int_1^2 (-1)^t e^{(-1)^t x}dt = \frac{1}{2\pi i}\int_{-1}^1 e^{xu}du = \frac{1}{i\pi x}\left(\frac{e^x - e^{-x}}{2}\right)$$

as required.

It is to be observed that the sum analog for cosh x goes through but

$$\frac{1}{2}\int_1^2 e^{(-1)^t x}dt = \frac{1}{2}\int_{-1}^1 u^2 e^{ux}dt = \frac{1}{2i\pi}\int_{-1}^1 ue^{xu}du = \left(\frac{x-1}{i\pi x}\right)\frac{\cosh x}{x}\cdot$$

Also solved by A. N. Aheart, M. Ash, José Asseo, D. W. Bailey, P. G. Bugl, M. M. Chowla (India), Craig Comstock, J. T. Darwin, Jr., D. E. Daykin, G. Di Antonio, A. D. Fine, C. L. Fountain, N. Ganesan (India), Mrs. A. C. Garstang, M. L. Glasser, Emil Grosswald, D. A. Hejhal, D. A. Horn, D. G. Kabe, D. B. Kirk, M. S. Klamkin, C. D. La Budde, E. S. Langford, Margaret M. La Salle, Stephen Libby, Beatriz Margolis, R. A. Moore, F. C. Smith, Al Somayajulu, M. R. Spiegel, Sidney Spital, L. A. Steen, C. M. Strauss, Maynard Tomer, H-O. Tønder (Denmark), C. J. Trauter (England), Gary Walls, A. Weinmann (England), K. L. Yocom, P. H. Young, and the proposer.

Partial Orderings from Total Orderings

5342 [1965, 1135]. *Proposed by R. B. Killgrove, the University of Wisconsin*

Suppose we have relations R_1, R_2, then we define $R=R_1\cap R_2$ as follows: xRy if and only if xR_1y and xR_2y. Now let S be the class of all total linear orderings of the positive integers, i.e., the relation L belongs to S if and only if (1) for any two positive integers x, y, we have xLy or yLx, (2) L is reflexive, (3) L is antisymmetric, (4) L is transitive. Now consider the relation D, "divides." Prove (i) D can be obtained in at least one way from intersecting certain members of S, and (ii) find the smallest (possibly an infinite cardinal) number of members of S needed to obtain D as their intersection. (Dedekind considered D as a partial ordering as an example of a lattice. Sometimes it is called Dedekind's ordering.)

I. *Solution by Seth Warner, Duke University.* For each $p\in P$, the set of primes, let $\leqq_p$ be defined as follows: if $a=p^ns$ and $b=p^mt$ where s and t are relatively prime to p, then $a\leqq_p b$ if and only if either $n<m$, or $n=m$ and $s\leqq t$. Clearly $\leqq_p$ is a total ordering (in fact, it is a well-ordering). It is also apparent that aDb if and only if for every $p\in P$, $a\leqq_p b$. Thus D is the intersection of the countable set of all $\leqq_p$, where $p\in P$.

To show that D is not the intersection of finitely many total orderings, let $L_1, \cdots, L_n$ be total orderings each containing D. As P is infinite, there exists $p\in P$ such that no $k\in[1, n]$ is p the last element of P for the total ordering L_k. Thus for each $k\in[1, n]$ there exists $q_k\in P$ distinct from p such that pL_kq_k. Let $q=q_1\cdots q_n$. Then pL_kq for all $k\in[1, n]$, since pL_kq_k and q_kL_kq. But p does not divide q. Therefore D is not the intersection of $L_1, \cdots, L_n$.

II. *Solution by Roy O. Davies, The University, Leicester, England.* The required cardinal is $\aleph_0$. Zorn's lemma easily yields:

(1) [Marczewski-Szpilrajn] Every partial ordering of any set E can be extended to a total ordering of E, in which the order of any one previously unordered pair may be pre-assigned. Hence (easily):

(2) [Dushnik and Miller] Every partial ordering of any set E is an intersection of total orderings, at most $(\text{card } E)^2$ being required, and this equals card E if it is infinite.

[If card E is finite or an aleph, the axiom of choice can be avoided in the proofs.]

(3) We complete the proof by showing that if D ("divides") is the intersection of a subclass T of S and k is any positive integer then card $T \geqq k$. Let p_i denote the ith prime number, and let $q_i = (p_1 p_2 \cdots p_k)/p_i$. Since not $p_i D q_i$, there is a member L_i of T such that $q_i L_i p_i$. For $i \neq j$ we have $p_j D q_i$, whence $p_j L_i q_i$ and by transitivity $p_j L_i p_i$; and similarly $p_i L_j p_j$; therefore by antisymmetry $L_i \neq L_j$. Thus we have found k distinct members of T.

Also solved by E. S. Langford, J. C. Morgan II, Donald Quiring, R. N. Raffelock, D. L. Silverman, and the proposer.

A Determinant of Binomial Coefficients

5343 [1965, 1135]. *Proposed by J. M. Gandhi, Maharani's College, Jaipur, India*

Show that

$\sum_{t=0}^{\gamma} \frac{\gamma!}{t!} = (-1)^\gamma D$, where D is the following determinant:

$$\begin{vmatrix} 0! & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1! & -\binom{1}{0} & \binom{1}{1} & 0 & \cdots & 0 & 0 \\ 2! & \binom{2}{0} & -\binom{2}{1} & \binom{2}{2} & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ (\gamma-1)! & (-1)^{\gamma-1}\binom{\gamma-1}{0} & \cdots & & & -\binom{\gamma-1}{\gamma-2} & \binom{\gamma-1}{\gamma-1} \\ \gamma! & (-1)^\gamma\binom{\gamma}{0} & (-1)^{\gamma-1}\binom{\gamma}{1} & & \cdots & \binom{\gamma}{\gamma-2} & \binom{\gamma}{\gamma-1} \end{vmatrix}$$

Solution by C. A. Church, Jr., West Virginia University. Replace γ by n. Move the first column to the last column; this eliminates the coefficient $(-1)^n$ on the right. Denote the entries by a_{ij} $(i, j = 0, 1, \cdots, n)$. Then, for $j = 0, 1, \cdots, n-1$,

$$a_{ij} = (-1)^{i-j}\binom{i}{j} = \begin{cases} 1, & i = j = 0, \\ 0, & i < j, \end{cases}$$

and $a_{in}=i!$. For $j=0, 1, \cdots, n-1$,

$$\sum_{i=0}^{n}\binom{n}{i}a_{ij}=\sum_{i=j}^{n}(-1)^{i-j}\binom{n}{i}\binom{i}{j}=\binom{n}{j}\sum_{i=0}^{n-j}(-1)^{i}\binom{n-j}{i}=0,$$

and

$$\sum_{i=0}^{n}\binom{n}{i}a_{in}=\sum_{i=0}^{n}\binom{n}{i}i!=\sum_{t=0}^{n}\frac{n!}{t!}\cdot$$

These elementary row operations reduce the last row to all zeros except the last element, but this is the value of the determinant (by expansion along the last row.)

If the same problem is considered with entries $b_{ij}=|a_{ij}|$, i.e., all entries nonnegative, then we have a representation for the familiar "rencontres" numbers $\sum_{t=0}^{n}(-1)^t n!/t!$.

Also solved by L. Carlitz, C. A. Church, Jr. (second solution), Eldon Hansen, Margaret M. LaSalle, F. D. Parker, D. P. Roselle, Sidney Spital, Selig Starr, J. H. van Lint, A. Weinmann (England), and the proposer.

Lineless Partition of a Finite Projective Plane

5345 [1965, 1136]. *Proposed by Carl Evans, Union College, Barbourville, Ky.*

Let Π be a finite projective plane of order $n>2$. Show that there is a partition of the points of Π into two sets such that neither set contains all of the $n+1$ points on a line.

Solution by Peter Yff, American University of Beirut, Lebanon. Take any triangle ABC in Π. Let one set (called S) consist of all points on AB and AC except B and C, plus another point P on BC. Then every line of Π, other than AB and AC, intersects S in at least one point but not more than three. Therefore neither S nor its complement contains an entire line.

When $n=2$, the statement is false. Since there are seven points, one of the sets contains at least four points. If it consists of four points of which no three are collinear, the complement is a line. If it contains five points, it must contain a line.

Also solved by Roy O. Davies (England), J. W. Di Paola, F. T. Howard, F. Kárteszi (Hungary), J. J. Martinez, J. G. Mauldon (England), F. D. Parker, William Smythe, Charles Wells, and the proposer.

The Coefficients in a Numerical Integration

5347 [1965, 1136]. *Proposed by Sidney Spital, Polytechnic State College, Pomona, California*

Show that the coefficients in Adams-Bashforth integration

$$\gamma_n = (-1)^n \int_0^1 \binom{-s}{n} ds$$

are given asymptotically by

$$\gamma_n \sim \frac{1}{\log n} + O\left(\frac{1}{\log^2 n}\right).$$

I. *Solution by W. F. Trench, Drexel Institute of Technology.* From Stirling's approximation,

$$(-1)^n \binom{-s}{n} = \binom{n+s-1}{n} = \frac{n^{s-1}}{\Gamma(s)} [1 + O(1/n)].$$

Hence, by integration by parts,

$$\gamma_n = \frac{1}{\log n} + \frac{1}{\log n} \int_0^1 \frac{\Gamma^1(s)}{\Gamma^2(s)} n^{s-1} ds + O(1/n).$$

There exists an upper bound for $|\Gamma'(s)/\Gamma^2(s)|$ in $[0, 1]$, and the integral on the right is less than $M/\log n$, which yields the result.

II. *Solution by A. Weinmann, Leicester, England.* This follows as a very special case of the asymptotic expansion of the generalized Bernoulli polynomials given by N. E. Nörlund, *Sur les valeurs asymptotiques des nombres et des polynômes de Bernoulli*, Rend. Circ. Mat. Palermo, (2), 10 (1961) 1–18. For, as is well known, $(-1)^n n! \gamma_n = B_n^{(n)}(0) = B_n^{(n)}$ (see, e.g., Nörlund, Section 1, or equation (38)), and the required asymptotic expansion can then be obtained from Nörlund's equations (21), or (31), or the equation after (41).

Also solved by L. Carlitz, M. L. Glasser, and the proposer.

Editorial Note. Using Haubel's integral for $1/\Gamma(s)$, Glasser derives the formula

$$\gamma_n = \frac{1}{\log n} + \frac{A}{\log^2 n} + O\left(\frac{1}{\log^3 n}\right),$$

$A = C^2 - \pi^2/6$, $C =$ Euler's constant.

The Sequence Function, $x_n = f(x_{n+1}, x_{n+2}, \cdots)$

5348 [1965, 1136]. *Proposed by Fred Galvin, St. Paul, Minnesota*

Show that there is a function f such that for any sequence $(x_1, x_2, \cdots)$ we have $x_n = f(x_{n+1}, x_{n+2}, \cdots)$ for all but finitely many n.

Solution by D. L. Silverman, Hughes Aircraft Company, Los Angeles. Define an *end-sequence* of $(x_1, x_2, \cdots)$ as a sequence of the form $(x_k, x_{k+1}, \cdots)$, where k is any positive integer. If $k > 1$, call such an end-sequence *proper*, and define the element x_{k-1} as the *predecessor* of $(x_k, x_{k+1}, \cdots)$ in the sequence $(x_1, x_2, \cdots)$. Well-order the sequences, and for each sequence s, let $f(s)$ be the predecessor of

s in the first sequence of the ordering of which s is a proper end-sequence. Then, for an arbitrary sequence $(x_1, x_2, \cdots)$, if m is the smallest subscript (if any) such that $(x_m, x_{m+1}, \cdots)$ is a proper end-sequence of a prior sequence in the ordering, $f(x_{n+1}, x_{n+2}, \cdots)=x_n$ holds for all n except $n=m-1$, and holds without restriction if $m=1$ or if no end-sequence of $(x_1, x_2, \cdots)$ is a proper end-sequence of a prior sequence in the ordering. This proves the stronger result that for any sequence $(x_1, x_2, \cdots)$, $x_n=f(x_{n+1}, x_{n+2}, \cdots)$ holds for all n with at most one possible exception.

Also solved by D. W. Bailey, R. D. Berlin, Roy O. Davies (England), S. J. Garland, Ellen Hertz, J. G. Mauldon (England), J. T. Rosenbaum, Leonard Shapiro, David Stanberry, L. A. Steen, B. L. D. Thorp, J. H. van Lint, and the proposer.

Editorial Note. All solutions to this problem used the axiom of choice in one way or another. Davies observed that the use of the axiom is to be expected because, using sequences of 0's and 1's, the result of the problem leads to the existence of a countable collection of sets $E_k \subset (0, 1)$, each with inner measure zero, and $\sum E_k = (0, 1)$.

A Kuratowski Closure and Complement Problem

5349 [1965, 1136]. *Proposed by S. P. Lloyd, Bell Telephone Laboratories, Murray Hill, N. J.*

With A a subset of a topological space X, let $\mathfrak{K}(A)$ be the smallest class of subsets of X with the properties (i) A is in the class, (ii) for every member E of the class, both E' and E^- are members of the class. It is known that $\mathfrak{K}(A)$ has at most 14 members and that there is a subset of the reals for which $\mathfrak{K}(A)$ does have 14 members (J. L. Kelley, *General Topology*, p. 57). If $\mathfrak{F}_1(A)$ is the field of subsets of X generated by the members of $\mathfrak{K}(A)$ show that $\mathfrak{F}_1(A)$ has at most 16,384 $(=2^{14})$ members, and give a subset of the reals for which $\mathfrak{F}_1(A)$ does have 16,384 members. Let $\mathfrak{F}_2(A)$ be the smallest field of subsets of X which has properties (i) and (ii) above. Give a subset of the closed unit interval for which $\mathfrak{F}_2(A)$ has infinitely many members.

Solution by J. C. Morgan II, University of California, Berkeley. We shall make use of the results of C. Kuratowski, *Sur l'opération $\overline{A}$ de l'Analysis Situs*, Fund. Math., 3 (1922) 182–199.

NOTATION. $a_1=A'^{-\prime}$, $a_2=A'^{-\prime-\prime-\prime}$, $a_3=A'^{-\prime-}$, $a_4=A^{-\prime-\prime}$, $a_5=A$, $a_6=A^{-\prime-\prime-}$, $a_7=A^-$, $a_1'=A'^-$, $a_2'=A'^{-\prime-\prime-}$, $a_3'=A'^{-\prime-\prime}$, $a_4'=A^{-\prime-}$, $a_5'=A'$, $a_6'=A^{-\prime-\prime-\prime}$, $a_7'=A^{-\prime}$ (cf. Kuratowski, p. 186). We shall identify intervals of the real line with the set of points of the intervals. Thus, $[a, b)=\{x \mid a\leqq x<b\}$, $(-\infty, a) =\{x \mid x<a\}$, etc. Let Q (I, resp.) denote the set of rational (resp., irrational) numbers, and let $(a, b)_Q=(a, b)\cap Q$, $[a, b]_I=[a, b]\cap I$, etc.

Part 1. The elements of $\mathfrak{F}_1(A)$ are the finite unions of elements of the form $p=\bigcap_{k=1}^7 \alpha_k$, where $\alpha_k=a_k$ or $\alpha_k=a_k'$ $(k=1, 2, \cdots, 7)$. Due to the relationship between the elements $a_1, a_2, \cdots, a_7$ (cf. Kuratowski, p. 186) it can be shown that there are at most 14 nonzero elements of the form p. Thus there are at most 2^{14} finite unions.

Part 2. A subset of the reals for which $\mathfrak{F}_1(A)$ has 2^{14} members is:

$$A = \left\{\frac{1}{n}\,\middle|\, n = 1, 2, \cdots\right\} \cup [2, 3) \cup (3, 4)_Q \cup [5, 6) \cup [6, 7]_Q \cup [8, 9]$$

$$\cup \left\{x \,\middle|\, 9 + \frac{1}{n+1} < x < 9 + \frac{1}{n}, \text{ for some } n \geqq 1\right\} \cup (10, 11) \cup \{12\}.$$

Part 3. Kuratowski gives a subset of the real line such that $\mathfrak{F}_2(A)$ is infinite (see loc. cit., p. 197; this example is also found in J. C. C. McKinsey & A. Tarski, *The Algebra of Topology*, Ann. Math., 45 (1944) 169). This example can be reproduced as a subset of the closed unit interval.

Verification of Part 1.

CASE 1: $\alpha_1 = a_1$. If $\alpha_k = a_k'$ (for $k > 1$) then $p \subseteq a_1 \cap a_k' \subseteq a_k \cap a_k' = 0$.
The only possibility is: $p_1 = a_1 \cap a_2 \cap a_3 \cap a_4 \cap a_5 \cap a_6 \cap a_7$.

CASE 2: $\alpha_1 = a_1'$, $\alpha_2 = a_2$. If $k \in \{3, 4, 6, 7\}$ and $\alpha_k = a_k'$ then $p \subseteq a_2 \cap a_k' \subseteq a_2 \cap a_2' = 0$.
The only possibilities are:
$p_2 = a_1' \cap a_2 \cap a_3 \cap a_4 \cap a_5 \cap a_6 \cap a_7$,
$p_3 = a_1' \cap a_2 \cap a_3 \cap a_4 \cap a_5' \cap a_6 \cap a_7$.

CASE 3: $\alpha_1 = a'$, $\alpha_2 = a_2'$, $\alpha_3 = a_3$. If $k \in \{6, 7\}$ and $\alpha_k = a_k'$ then $p \subseteq a_3 \cap a_k' \subseteq a_3 \cap a_3' = 0$.
The only possibilities are:
$p_4 = a_1' \cap a_2' \cap a_3 \cap a_4 \cap a_5 \cap a_6 \cap a_7$,
$p_5 = a_1' \cap a_2' \cap a_3 \cap a_4 \cap a_5' \cap a_6 \cap a_7$,
$p_6 = a_1' \cap a_2' \cap a_3 \cap a_4' \cap a_5 \cap a_6 \cap a_7$,
$p_7 = a_1' \cap a_2' \cap a_3 \cap a_4' \cap a_5' \cap a_6 \cap a_7$.

CASE 4: $\alpha_1 = a_1'$, $\alpha_2 = a_2'$, $\alpha_3 = a_3'$, $\alpha_4 = a_4$. As in case 3, we must have $\alpha_6 = a_6$, $\alpha_7 = a_7$.
The only possibilities are:
$p_8 = a_1' \cap a_2' \cap a_3' \cap a_4 \cap a_5 \cap a_6 \cap a_7$,
$p_9 = a_1' \cap a_2' \cap a_3' \cap a_4 \cap a_5' \cap a_6 \cap a_7$.

CASE 5: $\alpha_k = a_k'$ $(k = 1, 2, 3, 4)$, $\alpha_5 = a_5$. If $\alpha_7 = a_7'$, then $p \subseteq a_5 \cap a_7' \subseteq a_5 \cap a_5' = 0$.
The only possibilities are:
$p_{10} = a_1' \cap a_2' \cap a_3' \cap a_4' \cap a_5 \cap a_6 \cap a_7$,
$p_{11} = a_1' \cap a_2' \cap a_3' \cap a_4' \cap a_5 \cap a_6' \cap a_7$.

CASE 6: $\alpha_k = a_k$ $(k = 1, 2, 3, 4, 5)$, $\alpha_6 = a_6$. As in case 5, we must have $\alpha_7 = a_7$.
The only possibility is:
$p_{12} = a_1' \cap a_2' \cap a_3' \cap a_4' \cap a_5' \cap a_6 \cap a_7$.

CASE 7: $\alpha_k = a_k'$ $(k = 1, 2, 3, 4, 5, 6)$.
The only possibilities are:
$p_{13} = a_1' \cap a_2' \cap a_3' \cap a_4' \cap a_5' \cap a_6' \cap a_7$,
$p_{14} = a_1' \cap a_2' \cap a_3' \cap a_4' \cap a_5' \cap a_6' \cap a_7'$.

Thus, $\mathfrak{F}_1(A)$ has at most 14 atoms.

Verification of Part 2. Direct calculation establishes that

$p_1 = a_1$
$p_2 = a_1' \cap a_2 \cap a_5 = \{9\}$
$p_3 = a_2 \cap a_5' = \{9 + \frac{1}{n} \mid n = 1, 2, \cdots\}$
$p_4 = a_2' \cap a_3 \cap a_4 \cap a_5 = \{6\}$
$p_5 = a_2' \cap a_3 \cap a_4 \cap a_5' = \{3\}$
$p_6 = a_3 \cap a_4' \cap a_5 = \{2, 5, 8\}$
$p_7 = a_3 \cap a_4' \cap a_5' = \{11\}$
$p_8 = a_3' \cap a_4 \cap a_5 = (3, 4)_Q \cup (6, 7)_Q$
$p_9 = a_3' \cap a_4 \cap a_5' = (3, 4)_I \cup (6, 7)_I$
$p_{10} = a_3' \cap a_4' \cap a_5 \cap a_6 = \{7\}$
$p_{11} = a_5 \cap a_6' = \{\frac{1}{n} \mid n = 1, 2, \cdots\} \cup \{12\}$
$p_{12} = a_3' \cap a_4' \cap a_5' \cap a_6 = \{4\}$
$p_{13} = a_5' \cap a_6' \cap a_7 = \{0\}$
$p_{14} = a_7'$

Verification of Part 3. A countably infinite number of disjoint open subintervals of the unit interval is designated. There is a one-to-one correspondence I between this family of subintervals and the ordinal numbers $\alpha < \omega^\omega$. If $\alpha < \omega^\omega$ then we can find a subset of $I^{-1}(\alpha)$ of order type α.

Also solved by Donald Quiring, J. G. Mauldon (England), and the proposer.

RECENT PUBLICATIONS AND PRESENTATIONS

EDITED BY R. A. ROSENBAUM, Wesleyan University
COLLABORATING EDITORS: KENNETH O. MAY, University of Toronto and
E. P. VANCE, Oberlin College

Materials intended for review should be sent to Prof. Kenneth O. May, American Mathematical Monthly, Department of Mathematics, University of Toronto, Toronto 5, Ont., Canada.

Variational Methods for the Study of Nonlinear Operators. By M. M. Vainberg. With a chapter on Newton's method, by L. M. Kantorovich and G. P. Akilov. Translated and supplemented by Amiel Feinstein. Holden-Day, San Francisco, London, Amsterdam, 1964. x+323 pp. $12.95.

The so-called variational method for solving an equation

$$F(x) = 0 \tag{1}$$

or for treating an eigenvalue problem

$$F(x) = \lambda x \tag{2}$$

may be described as follows: if $f(x)$ is a real valued function defined in a domain D then a necessary condition for $f(x)$ to take an extremum in an interior point of D is that at that point

$$df = 0 \text{ or, equivalently, } \operatorname{grad} f = 0. \tag{3}$$

This is no longer true if the extremum is "conditional"; if, e.g., x is subject to the condition that it lie on the surface of a sphere, then (3) is replaced by

$$\operatorname{grad} f = \lambda x. \tag{4}$$

Consequently if the given F in (1) or (2) is a "potential operator," i.e., if there exists a real valued function f such that

$$F = df, \quad \text{or} \quad F = \operatorname{grad} f, \tag{5}$$

then any x_0 giving an extremum or conditional extremum as described above will be a solution of (1) or (2) respectively.

This method of reducing problems of the form (1) and (2) is elementary if D is in a finite dimensional Euclidean space. It is classical in certain function spaces if F is linear (i.e., if f is quadratic).

Since about 1930 the method has been extended to nonlinear operators F in function spaces. It is the object of the book under review to give an up-to-1955 account of such investigations, many of which are due to the book's author.

To carry through as the author does the above program in Banach or Hilbert space one obviously needs a treatment of differentials (Gateaux and Fréchet) in such spaces and of various forms of continuity. Such treatment is contained in chapter I. Since the variational method is applicable only if F is a potential operator, i.e., if (5) is true, chapter II is devoted to properties of potential operators and to the formulation of sufficient conditions for F to be a potential operator. Here Stieltjes integrals in a Banach space come into play. A treatment of such integrals is therefore incorporated in chapter I.

Chapter III deals with critical points of f (i.e., points where (3) is satisfied) and contains conditions sufficient for the existence of an extremum of f in a ball in a reflexive Banach space. Here the "weak" compactness of such a ball and the "weak" continuity of $f(x)$ play a decisive role. A sufficient condition for the latter property is the compactness of the operator $F = \operatorname{grad} f$.

Chapter IV deals with conditional extrema and conditionally critical points leading to a type of equation of which (4) is a special case. In particular the case in which the point x lies on a hyperboloid or an ellipsoid in a Hilbert space is treated. In case the ellipsoid is a sphere in a separable Hilbert space, a proof for the existence of infinitely many conditionally critical points is given if the scalar f is even. The proof uses the Lusternik-Schnirelman "category" theory of which a short exposition, mainly without proofs, is given.

The first part of chapter V deals with eigenelements of a nonlinear operator F. An element $x_0 \neq 0$ is called an eigenelement of F if (2) is satisfied provided that $F(0) = 0$, a definition which agrees with the usual one for a linear F. Thus the eigenvalue problem is closely connected with a suitable "conditional" problem, and the results of chapter IV play a decisive role. Particular attention is given to the eigenvalue problem in Hilbert space for operators of the form $F = AF_1$ where A is linear and self adjoint while F_1 is a potential operator. (Operators of this form have been already considered in chapter III.)

The second part of chapter V deals with the existence of "bifurcation points" of (2), again under the assumption that $F(0) = 0$, which has the consequence that the point $(x = 0, \lambda)$ is a solution of (2) for all λ. A point $(x = 0, \lambda = \lambda_0)$ is

then called a bifurcation point for (2) (or for F) if every neighborhood of this point contains a solution (x, λ) of (2) with $x \neq 0$. In particular the connection between the bifurcation problem for F and the eigenvalue problem for its differential dF is discussed.

Chapter VI deals with operators and functionals of the special type needed for the applications to nonlinear integral equations given in the following chapter. For instance in the Hammerstein equation the operator Γ of the form

$$\Gamma(u) = \int K(x, y) g(u(y), y) dy$$

appears. This is the product of the Nemytsky operator $u \to g(u(y), y)$ with the linear operator

$$Av = \int K(x, y)\, v(y) dy. \tag{6}$$

Consequently these two operators are treated in detail, particularly concerning their differentiability and the property of having a potential. The underlying space is usually L^p with $p \geqq 1$. Similar investigations are made concerning the Liapunov-Lichtenstein operator, which consists of an infinite sum of integrals of the form

$$\int_0^1 \cdots \int_0^1 K(s, t_1, t_2, \cdots t_n) x(t_1) x(t_2) \cdots x(t_n) dt_1, dt_2, \cdots dt_n.$$

The remainder of this chapter is devoted to the eigenvalue and expansion problem concerning the linear operator A defined in (6) and its square root in case A is symmetric and completely continuous. It is classical that these problems can be treated by considering certain conditional extremum problems for a quadratic form associated with A if the underlying space is L^2. It is shown that this method may be extended to the case where A maps L^q into $L^p((1/p)+(1/q)=1)$ with $q<2$.

Chapter VII, the last of Vainberg's book, applies the preceding theory to existence, uniqueness, eigenvalue and bifurcation problems of nonlinear integral equations, mainly of the Hammerstein type (or systems of such) under various assumptions.

The last chapter of the present book deals with the generalization to Banach spaces of Newton's well-known method for solving equations. It is a translation of the last chapter of the book "Functional Analysis and Normed Spaces," by G. P. Akilov and L. V. Kantorovich. This book has since been translated in its entirety (1964).

As to prerequisites for reading his book, Professor Vainberg states that he assumes familiarity with functional analysis to the extent of the first five chapters of Lusternik-Sobolev's book, Elements of Functional Analysis (1951),

(translated 1961, Ungar, New York). However, the translator added a supplement containing a brief survey of results of Banach space theory used in the text.

Throughout the book proofs are based on properties of the function space involved and not as is often the case in the older literature on "going to the limit" with finite dimensional subspaces. While the latter method is still valuable for approximation purposes, the former is in keeping with the modern trend and seems to be more systematic from the theoretical point of view.

As to the notion of weak continuity, the author uses throughout the traditional "sequential" definition in contradistinction to the one based on the weak topology (mentioned in passing p. 75/76). This reviewer wonders whether the use of the latter definition might not result in a still more systematic treatment since the general theorems on compact topological spaces would be available. In many of the spaces occurring in the book the two definitions are equivalent.

All in all this is a very informative book which will be valuable to all mathematicians interested in the subject matter treated.

The practical value is enhanced by an extensive bibliography.

E. H. ROTHE, University of Michigan
and Western Michigan University

Calculus of Several Variables. By Casper Goffman. Harper and Row, New York, 1965. ix+182 pp. $7.00.

The chapter headings in this concisely written book are as follows: I Euclidean Space, II Mappings and their Differentials, III Mappings into the Reals, IV Main Theorems on Mappings, V Manifolds—Differential Forms, VI Vector Analysis. There are about 250 exercises (many calling for proofs which are not in the text) and an abbreviated, although possibly adequate index. The section on integration of real valued functions on $E^{(n)}$ is by far the most satisfying. The key to the arguments is the observation that any open set in $E^{(n)}$ is the union of countably many closed n-dimensional cubes which are pairwise non-overlapping in the sense that $I \cap J$ contains no interior points of the n-dimensional cubes I or J. The climax of this section is the theorem that a bounded function on an n-dimensional cube is integrable if and only if its points of discontinuity comprise a set which is the union of countably many sets of zero volume. I also found the method of presentation of definitions to be literarily more satisfying than the usual method. At the same time, there are precisely 7 figures, some of which illustrate trivial matters, and there is no motivation of the concepts whatsoever. Conceivably, the motivation is to be found in previous courses (this is obviously a textbook) but, for example, we get no more information than "Laplace's equation $\Delta f = 0$ arises in many places in pure and applied mathematics." Only one major class of problems is mentioned—problems of diffeomorphism. One does not get the slightest hint of differential equations on manifolds or of why one would want to integrate differentiable forms. There is not even a schematic diagram of a tangent vector. Moreover, I do not believe that

the text convinces one of any advantage of differential forms and exterior algebra. Aside from minor errors, I was unable to find a definition of the notation (page 154)

$$\int B_n \cdots dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

in previous chapters. The proof of the formula for differentiation under the integral sign depends on the mean value theorem without showing that the mean value theorem yields an integrable function. Finally, the key lemma in the proof of Stokes theorem (page 158) is not clearly used in the body of the proof, although a student familiar with certain techniques should be able to fix up the latter proof.

MALCOLM GOLDMAN, New York University

Linear Approximation. By Arthur Sard. Mathematical Surveys, No. 9, Amer. Math. Soc., 1963. xi+544 pp. \$16.80.

In this book, Professor Sard conducts us on a guided tour of one of the larger edifices which comprise approximation theory, viz., the theory of approximating one linear operator or functional by another. The author has been the principal architect of this structure during the past few years of vigorous development.

The central problem of the book is this: one is presented with a linear operator L and with a family $\mathfrak{F}$ of linear operators; an element F of $\mathfrak{F}$ is sought for which $\|L-F\|$ is a minimum. Unfortunately, this simple description does not convey any hint of the great diversity of interesting particular problems which can be couched in these terms. Actually, many problems of numerical analysis, such as the smoothing of data, numerical integration, interpolation, and the economical representation of common functions are subsumed by this general question. As an illustration, one may cite the typical problem of numerical integration in which it is required to approximate the linear functional $Lx=\int_{-1}^{1} x(t)dt$ by a linear functional of the form $Fx=\alpha x(-1)+\beta x(0)+\gamma x(1)$. After deciding on an appropriate normed linear space as the situs for x, one may ask for an optimum determination of α, β, and γ; i.e., a choice making $\|L-F\|$ a minimum. The author devotes 15 pages to the analysis of this very problem in a variety of practical situations. General numerical quadrature and interpolation formulas in one variable are the subject of the lengthy Chapter II (81 pages). The same degree of comprehensive scholarship characterizes the treatment of each topic taken up by the author and accounts for the book's size.

The theoretical aspects of the subject depend heavily upon representation theorems for linear functionals and operators in the classical spaces and in various other spaces introduced by the author because of their usefulness in numerical analysis. This book will be a convenient source of information on this topic as well as on its central topic of optimum formulas.

E. W. CHENEY, University of Texas

Continuous Mappings and Conditions of Monogeneity. By Yu. Yu. Trokhimchuk. Israel Program for Scientific Translation. Transl. R. Mandl. Daniel Davey, New York, 1965. vi+133 pp. $10.25.

This book is the translation of a revised version of the author's 1960 dissertation. Based on methods of analytic topology and the theory of real variables (perfect sets, category, existence of differentials, etc.) it presents generalizations of the conditions of Looman-Menchov and of Bohr that a continuous function of a complex variable be analytic in a region. The material is accessible to second year graduate students. Some open problems, mostly relating to the existence of counterexamples, are pointed out.

The translation is quite good; there seem to be few misprints. The translator points out some inaccuracies in the original text. He does not, however, warn the reader that in the author's terminology a set is of "second category" just in case its complement is of the first category.

E. C. SCHLESINGER, Connecticut College

Elements of the Theory of Probability. By Émile Borel. Translated from the French by John E. Freund. Prentice-Hall, Englewood Cliffs, N. J., 1965. 178 pp. $6.50.

The book under review is an English translation of Émile Borel's *Éléments de la théorie des probabilités*. As stated in the Preface, this book may be considered as an introduction to the distinguished author's classic treatise *Traité du calcul des probabilités et de ses applications*. The chapter headings give an indication of the wide variety of interesting topics discussed in this volume: 1. The game of heads or tails; 2. Definitions and theorems; 3. Approximation formulas; 4. Further study of the game of heads or tails; 5. The law of large numbers; 6. The law of chance; 7. Definition of geometrical probability; 8. Some problems of geometrical probabilities; 9. The introduction of arbitrary functions; 10. Errors of measurement; 11. The problem of choice in the continuum; 12. The discrete case; 13. Statistical problems; 14. The continuous case; 15. The determination of causes. There are also several appendices; of particular interest are those dealing with psychological games and the imitation of chance, objective and subjective probabilities, and the Petersburg paradox.

Although this volume presents a very lucid and informal introduction to many important and controversial problems associated with probability theory, it is not suitable as a text for what might be called a conventional course in probability theory. In the opinion of this reviewer, however, this book should be read by students of probability theory and mathematical statistics as well as by students of philosophy and methodology of science.

A. T. BHARUCHA-REID, Wayne State University

Numerical Solution of Partial Differential Equations. By G. D. Smith. Oxford, London, 1965. 188 pp. $4.00.

This book is an introduction both to partial differential equations and to their numerical solution. It begins with a description of the methods used to solve the heat equation. The concepts of stability and convergence are introduced in the next chapter, where some proofs are given. The final two chapters are devoted to hyperbolic and elliptic equations. The chapter on hyperbolic equations is not restricted to the wave equation; it even discusses shocks.

The aim of the author was to describe the state of affairs honestly, giving proofs when they can be understood by a wide audience, and citing the literature otherwise. He has succeeded in this because he frequently gives proofs of special cases of the theorems he states.

The book contains many examples and exercises, and it also contains solutions to the exercises.

G. W. HEDSTROM, University of Michigan

The Collected Papers of Emil Artin. Edited by Serge Lang and John T. Tate. Addison-Wesley, Reading, Mass., 1965. xvi+560 pp. $13.50.

This volume reproduces all the papers of Emil Artin (1898–1962). They are arranged by topic and appear in chronological order within each topic heading. A tripartite preface by the editors contains a short outline of Artin's career, a brief but intense appreciation of Artin the teacher, and a concise discussion of some of Artin's more important conjectures, their motivation, and their current status. The table of contents, which doubles as a bibliography, is followed by a list of Artin's books and lecture notes, none of which is reprinted in the present work. This book is a fitting tribute to a great mathematician by two of his best students; it deserves to be read and recommended by all who would keep his memory alive.

F. E. J. LINTON, Wesleyan University

Numerical Analysis. By I. M. Khabaza. Pergamon Press, Oxford and New York, 1965. 242 pp. $5.00.

This book has been written for the undergraduate student interested in numerical methods associated with digital computers. The author has written a clear book, with pertinent examples to clarify the text and an extensive list of problems to exercise the student. Although this is an elementary book, the author constantly directs the reader's attention to errors in these numerical methods, and he shows how these errors adversely affect the solution.

The book's contents by chapter are: 1. Digital Computers. 2. Desk Machines. Errors in Computation. 3. Finite Difference Methods. 4. Recurrence Relations and Algebraic Equations. 5. Numerical Solution of Ordinary Differential Equations. 6. Matrices. 7. Relaxation Methods. 8. Numerical Methods for Unequal Intervals.

LEON LEVINE, Scientific Data Systems

Elements of the Theory of Elliptic and Associated Functions with Applications. By Mahadur Dutta and Lokenath Debrath. World Press, Calcutta, 1965. 290 pp. Rs. 22.50, 35/-.

This book presents an account of elliptic functions and modular functions covering roughly the same material as E. T. Whittaker and G. N. Watson, "A Course in Modern Analysis," and using the same method of approach. A few topics which are not treated in Whittaker and Watson are mentioned: Riemann surfaces, algebraic curves, uniformisation, solution of the quintic. These subjects are treated so sketchily that it would have been better not to include them at all. There are several inaccuracies, in particular the alleged theorem that two simply periodic functions with the same periods, the same zeros and the same poles must be linearly dependent is not true, as the example e^z, $\exp\{e^z\}$ shows.

W. H. J. FUCHS, Cornell University

The Foundations of Astrodynamics. By Archie E. Roy. Macmillan, New York, 1965. xiv+385 pp. $10.95.

The advent of the "space age" has caused a remarkable rebirth of interest in and a need for textbooks in celestial mechanics. Professor Roy's book is one of the better products of this resurgence. Following a brief synopsis of the principal features of the solar system and a discussion of coordinates and time he presents the two, three and n-body problems of central force motion. Spheres of influence of planetary masses are discussed as well as Tisserand's criterion for the identity of comets. In general the treatment is lucid. There are quite a few errors in proof reading which may cause the uninitiated student to think hard. The student reading the book would gain a great deal if there were more illustrative examples clearly worked out.

Perturbation theory is treated by the standard variation of parameters technique. A brief discussion of Lagrange's equations and the Hamilton-Jacobi theory introduces the student to these more elaborate techniques. The earth-moon system is considered in considerable detail.

By far the most valuable contributions made by the author are his discussions of rocket and artificial satellite motions. The effects of atmospheric drag, earth's oblateness, and so forth are well described. Transfer orbits for planetary probes, escape orbits from a planet or the moon, interplanetary navigation and the effects of errors in interplanetary orbits are among the topics discussed in detail. There are brief discussions but no numerical examples of methods for determining orbits from observation.

The book concludes with five appendices giving astronomical constants and numerical data concerning the solar system. Answers are given for the numerous problems with which each chapter concludes. The format of the book is excellent. It is up-to-date and should arouse the interest of the student by the discussions of currently provocative problems of space flight.

S. W. MCCUSKEY, Case Inst. of Tech.

Differential Equations. By Shepley Ross. Blaisdell, Waltham, Mass., 1964. 608 pp. $10.50. (A shorter version, essentially the same as the first nine chapters, is available under the title "Introduction to Ordinary Differential Equations." Blaisdell, 1966. viii+337 pp. $7.50.)

The book is divided into two distinct parts. The first part introduces the elementary methods for solving ordinary differential equations and considers simple applications to spring problems and simple electrical circuits. The treatment of power series techniques is perfunctory and stops short of any real applicability.

In spirit, the first chapter in the second part belongs in the first part, providing a superficial study of approximation and numerical methods. The next three chapters are more theoretical and give a good introduction to the theory of the Laplace transformation and the basic theorems on existence and uniqueness of solutions for first order equations and for linear equations. Fourier series are introduced through the study of Sturm-Liouville systems. Theorems whose proofs are considered difficult are clearly stated without proof.

The most interesting material in the book is contained in the chapter on nonlinear equations. The author discusses the phase plane and critical points with applications to linear systems and to non-linear conservative systems. He then motivates and discusses the Poincaré-Bendixson theorem, the theorem of Lienard-Levinson-Smith, and the method of Kryloff and Bogoliuboff, giving applications of these theorems in several special cases.

P. E. BEDIENT, Franklin and Marshall College and The University of Michigan

The Theory of Matrices in Numerical Analysis. By Alston S. Householder. Blaisdell, Waltham, Mass., 1964. 272 pp. $9.50.

This book is an excellent and scholarly approach to the role of matrices in numerical analysis. Classical algorithms for reduction of matrices to tridiagonal form, solution of linear systems, inversion, and other standard matrix problems are presented here, not as separate and distinct methods, but as part of a unified theory. A comprehensive discussion of norms and their applications to localization of characteristic roots, convergence criteria and error analysis is given. In addition Dr. Householder has simplified and enriched the content with many original ideas.

The book provides a solid background for the numerical analyst and for anyone interested in the practical applications of matrix theory, but it is far more than just a list or description of numerical methods. It is an excellent reference book for anyone doing research work in matrix theory. The problems and exercises alone give a varied and extensive sampling of recent results in this field. In this connection it would be helpful to have the references numbered, and a more complete list of the notation used is also needed; but these are very minor criticisms of a major book, which will undoubtedly become a classic in its field.

E. V. HAYNSWORTH, Auburn University

Vector and Tensor Analysis. By K. Pach and T. Frey. Akadémiai Kiadó, Budapest, 1964. 595 pp. $14.00.

The volume under review is an exposition of the methods and principal results of vector analysis as well as of some elementary tensor analysis. Written for engineers, there is a strong emphasis on the geometric, intuitive meaning of the quantities which are discussed, and on applications of the theory. The authors also pay more attention than is usual in such texts to a careful development of the statement and proof of the results achieved.

Arranged in a systematic, methodical sequence, the topics under discussion include vector algebra, vector functions of a scalar variable, scalar functions of a vector variable, vector fields, second order tensor algebra and tensor analysis. As outgrowths of these subjects, the authors treat such customary matters as the gradient, divergence, curl, the integral theorems of Gauss and Stokes. They also discuss in considerable detail some less usual topics including eigenvalue theory, non-stationary vector fields, Green's theorems and potential theory.

With the exception of a short final chapter, dealing briefly with multidimensional spaces and curvilinear coordinates, the underlying space of the entire volume is a three dimensional Euclidean space having a fixed Cartesian system of coordinates. The authors define vectors entirely by their geometric properties. Second and third order tensors are defined in terms of vectors by what is, in effect, the quotient rule. Strict adherence to this point of view (except in the final chapter) has severely circumscribed the scope of the tensor theory included in the book. Tensors of order $n(n>3)$ are not defined. The transformation laws for vector and tensor components and their invariance properties under changes of the coordinate system are not developed. Concepts which arise in curved Riemann spaces simply do not appear. There is no mention of covariant differentiation, parallel displacement, or the Riemann curvature tensor.

In its purely vectorial aspects, however, the volume has some clear advantages. An outstanding feature of the book is the very large number of applications to geometry and physics. These include many aspects of geodesics, particle dynamics, hydrodynamics, potential theory, elasticity and electromagnetic theory. The style of the book is leisurely, with a profusion of illustrations. Although helpful to the beginner, the lengthy discussion may generate some impatience in an experienced reader. The text proceeds in a logical development with considerable clarity of exposition. Many excellent exercises for the student are included.

AARON FIALKOW, Polytechnic Institute of Brooklyn

Handbook of Mathematical Functions. Edited by Milton Abramowitz and Irene A. Stegun. Dover, New York, 1965. 1043 pp. $4.00.

In the age of computers what is the place of the book of tables? In 1954 a conference considered this question and rightly concluded that with a general increase in computation, tables would become more, not less, important. The outcome of the conference was that the National Science Foundation and the

National Bureau of Standards set up a distinguished committee to advise on the preparation of a modern handbook of tables. Under the editorship of M. Abramowitz and Irene A. Stegun the book was divided into nearly thirty sections, and each was handled by members of the staff of the Mathematics Division of the National Bureau of Standards.

It would take too much space to list the sections; certainly it is more comprehensive than any previous general handbook. Each section has a well-written introduction giving extensive formulae and many well-chosen references. The number of significant figures in the tables varies from 18 for the exponential function to 5 for the parabolic cylinder functions. Generally, this kind of choice seems to have been made in accord with practical experience.

The printing and layout are extremely good and the book is pleasant to handle, although the publishers might consider issuing individual sections for classroom use. The price needs special note: four dollars for a book of over 1000 pages. This handbook will assuredly become a classic.

J. J. FLORENTIN, Brown University

Funktionalanalysis und numerische Mathematik. By L. Collatz. Springer-Verlag, Berlin, 1964. xvi+371 pp. 58 DM.

This is a comprehensive treatment of the two topics, iterative methods and monotone operators, on which the author and many of his students have concentrated in recent years. Of course the book includes more—an exposition of the required functional analysis, for example, and a useful discussion of Tchebycheff approximation and the exchange algorithm—but the spotlight is on those first topics.

The fundamental theorem on iterative methods is the contraction mapping principle: the sequence defined by $u_{n+1}=T(u_n)$ converges to the (unique) solution of $u=T(u)$, given that T maps a complete metric space into itself with $\rho(T(v), T(w)) \leqq P\rho(v, w)$, for some $P<1$ and all v, w. The author first generalizes this result, following Schröder, and then later specializes it to linear problems. In this case $T(u)=Au+b$, and the spectral radius of the linear operator A governs convergence; perhaps this should have been proved in infinite dimensions, but the finite-dimensional case leads to a compact discussion of overrelaxation and alternating direction methods.

Monotone operators generally arise in the presence of maximum principles, which are an essential feature of many elliptic and parabolic problems. Like self-adjointness and positive definiteness, this feature ought to be preserved numerically; it might even have been possible to give a general recipe for producing monotone operators in a class of elliptic problems. The author devotes his attention to working out several examples in detail, emphasizing the relevance of his theory to nonlinear problems.

GILBERT STRANG, Massachusetts Institute of Technology

Applications of Graph Theory to Group Structure. By Claude Flament. Prentice-Hall, Englewood Cliffs, N. J., 1963. 142 pp. $6.95.

The word "group" in the title means organization, and the intent of this monograph is to show that concepts from graph theory are relevant to the formal study of group behavior. Chapter 1 has introductory material (sets, ordered sets, connectivity, etc.). Chapter 2 discusses organization of a group and its communication channels to perform a task. Chapter 3 discusses models of positive and negative relationship, and their adjustments (balancing processes). Some interesting theorems are given here, but the author has missed the easy proof of the "conjecture" on page 107. There is a substantial bibliography.

The heart of the book is contained in the concluding sections of Chapters 2 and 3 (Applications), which unfortunately occupy only thirteen pages. The examples cited there are enough to stimulate interest, which was perhaps the extent of the author's ambition, but this reader was left with the feeling that the feast terminated with the appetizer.

A. J. HOFFMAN, Thomas J. Watson Research Center

Albert Einstein and the Cosmic World Order. By Cornelius Lanczos. Interscience, New York, 1965. 139 pp. $3.95.

This set of six lectures delivered at the University of Michigan in 1962 is an exposition for the educated layman of the fundamental ideas of Einstein's special and general relativity theory. It is a more serious, adult treatment than most "popular" writings about relativity. Rather than considering specific applications, Lanczos emphasizes basic concepts such as reference frames and covariance, Minkowski space-time (including an unusually lucid treatment of the twin paradox), basic ideas of geometry from Euclid to Gauss and Riemann, and the reasons that drove Einstein to a geometrical description of physics. These basic concepts are very clearly explained, using hardly any equations; interesting insights into Einstein's personality and into the history of physics of his time are found throughout the book. This book would be particularly appropriate for non-specialists who would like to appreciate the greatness of Albert Einstein by exploring the fascinating world order brought about by his relativity theory.

DIETER BRILL, Yale University

Elements of Modern Algebra. By Sze-Tsen Hu. Holden-Day, San Francisco, 1965. x+208 pp. $8.95.

This book is designed as a text for upper level undergraduates and beginning graduate students. The seven chapters are: "Sets, Functions and Relations," "Semigroups," "Groups," "Abelian Groups," "Rings; Integral Domains and Fields," "Modules, Vector Spaces and Algebras," and "Categories and Functors." The book is carefully written; proofs are generally very detailed and easy to follow. Exercises in the early part of the book tend to be routine and easy

but the later chapters have some exercises which will challenge good students.

The author's homological bent is revealed throughout in his choice of notation, of terminology and of subject matter. In the treatment of groups, for instance, the emphasis is decidedly on abelian groups; very little space is given to permutation groups and none at all to the linear groups, but exact sequences and commutative diagrams abound.

The book's approach to universal structures, such as free groups, free abelian groups, and tensor products, is superior to the usual ones. Its repeated reliance on the universal mapping property that characterizes such structures is specially effective.

All in all, the student will find here many important ideas and techniques that will not only serve him well in later studies in algebra but will be useful in other areas of mathematics.

In the reviewer's opinion a serious defect of the book as a text is that it gives a broad collection of definitions, concepts, lemmas and theorems, but in general gives little indication as to why these are important, either by use of examples, or applications in later theorems. Too often the student will have to wait until a later course to realize any payoff for his efforts. For a specific example, a detailed proof of "The Four Lemma" for exact sequences is given in Chapter III, but the reviewer was unable to find any further mention of this proposition in any later chapter. The last chapter, "Categories and Functors," is twelve pages long and essentially consists of definitions, a few immediate consequences of the definitions, and a list of examples. One wonders if, for the student at this level, there might not be other topics to which he could more profitably devote his time.

A. C. MEWBORN, University of North Carolina

Spanish-English/English-Spanish Mathematics Dictionary. By Mariano García. Hobbs-Dorman, New York, 1965. vii+78 pp. $2.50.

In this day of rapidly growing interest and participation in contemporary mathematical developments on the part of the Spanish-speaking world of Latin America, the appearance of this first Spanish-English/English-Spanish mathematics dictionary is particularly timely.

Professor García's dictionary is compact rather than compendious, being confined to around 3000 of the most commonly used mathematical terms and phrases. The words and phrases selected for inclusion were arrived at after comparisons of comparable foreign-English mathematics dictionaries for several other languages. Such refinements as indications of parts of speech, gender of nouns, and cross-references were omitted in the interest of more rapid publication. These may appear in a later edition. The printing, paper, and binding are of good quality, and the fact that the present paper-back edition is inexpensive should contribute to its wide use.

TRUMAN BOTTS, University of Virginia

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo (University of Buffalo), Buffalo, New York 14214. Items must be submitted at least two months before publication can take place.

PERSONAL ITEMS

Professor R. H. Bing, University of Wisconsin, has been appointed to the President's Committee on the National Medal of Science.

Concordia College: Professor W. L. Miser, Sweet Briar College, has been appointed Visiting Professor; Assistant Professor Charles Heuer, University of Missouri, has been appointed Assistant Professor; Assistant Professor Arne Garness has been named Acting Chairman of the Department of Mathematics.

University of Georgia: Professor George Adomian, Pennsylvania State University, has been appointed Barrow Professor of Mathematics; Assistant Professor J. T. Hardy, University of Southwestern Louisiana, and Dr. R. B. Sher, University of Utah, have been appointed Assistant Professors; Associate Professor J. G. Horne, Jr. has been promoted to Professor; Associate Professor J. W. Jewett has been appointed Executive Secretary of the Survey Committee of the Conference Board of the Mathematical Sciences and will be on leave for 1966–67.

Mr. Roger Backen, Educational Services of California, Oakland, California, has been appointed Chairman of Mathematics with the International School Association, Bangkok, Thailand.

Mr. R. M. Gordon, Scientific Data Systems of Santa Monica, has been appointed Assistant Director of the Computer Facility at the University of California, Irvine.

Mr. H. H. Harman, System Development Corporation, Santa Monica, California, has been appointed Senior Research Psychologist at Educational Testing Service, Princeton, New Jersey.

Assistant Professor K. O. Leland, University of Virginia, has been appointed Assistant Professor at the Illinois Institute of Technology.

Professor L. E. Mehlenbacher, University of Detroit, has been appointed Associate Dean of the Graduate School.

Professor M. W. Oliphant, Georgetown University, has been appointed Dean of Academic Affairs at Hawaii Loa College.

Associate Professor F. G. Scorsone, Eastern Kentucky University, has been promoted to Professor.

Assistant Professor E. M. Stone, California State College, Long Beach, has been appointed Assistant Professor at State University College of New York at Oneonta.

Assistant Professor L. G. Black, Purdue University, died on May 25, 1966. He was a member of the Association for 37 years.

Professor Emeritus L. M. Coffin, Coe College, died on April 14, 1966. He was a charter member of the Association.

Mr. Marcel Delcourte, Les Assurances du Boerenbond Belge, Louvain, Belgium, died on June 18, 1966. He was a member of the Association for 11 years.

Professor E. D. Rainville, University of Michigan, died on April 29, 1966. He was a member of the Association for 38 years.

Professor Emeritus L. J. Reed, Johns Hopkins University, died on April 28, 1966. He was a charter member of the Association.

JOHN VON NEUMANN CHAIR OF APPLIED MATHEMATICS

The University of Brussels has created the John von Neumann Chair of Applied Mathematics under the sponsorship of IBM. The first lecturer appointed to the chair is Professor E. Stiefel of Zürich.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL MEETING OF THE IOWA SECTION

The 53rd regular meeting of the Iowa Section of the MAA was held at the Central College, Pella, on April 15, 1966. Chairman R. V. Hogg presided. Total attendance was 105, including 56 members of the Association. Routine business was considered during the afternoon meeting.

A treasurer's report was given and a balance of $264.32 was indicated.

The following officers were elected: Chairman: W. L. Waltmann, Wartburg College; Vice-Chairman: Rev. J. L. Friedell, Loras College; Secretary-Treasurer: E. L. Canfield, Drake University; Vice-Secretary-Treasurer: B. E. Gillam, Drake University.

The following papers completed the program:

1. *Quasi-algebraic functions*, by E. S. Allen, Iowa State University.

Antonio Salmeri denotes as "quasi-algebraic" those functions of x and $[x]$ which are algebraic in x. (Gior. di Mat., vol. 91, 1963.) Those functions described by him, and those considered in this report, are algebraic in $[x]$ also. Solutions of linear equations are first considered. The paper then considers the possibility of replacing the real number x by the complex number $x+yi$.

2. *On order convergence in partially ordered sets*, by R. F. Anderson, Iowa State University, introduced by the Chairman.

3. *Generalized derivatives and integrals*, by D. L. Hansen, Westmar College, Le Mars.

Definitions of a generalized derivative have been given in the following three papers: K. Menger, this MONTHLY, October, 1957, pp. 58–70; K. Menger and S. Shü, *Proc. Nat. Acad. Sci.*, 1955, pp. 591–595; G. Pall, this MONTHLY, October, 1957, pp. 71–78. It is shown that the given definitions are equivalent. A definition is given for a generalized integral (using Pall's notation). Various analogs of integration theorems from complex analysis are then explored.

4. *Approximation of real continuous functions on the real line by infinitely differentiable functions*, by L. E. Pursell, Grinnell College.

Most extensions of the Weierstrass Approximation Theorem apply only to bounded functions (See: M. H. Stone, A Generalized Weierstrass Approximation Theorem, *MAA Studies*, vol. 1, 1962). It is shown here that any real continuous function on the real line R can be uniformly approximated on R by infinitely differentiable functions. It follows that if the set $C^\infty(R)$ of all real infinitely differentiable functions on R is considered as an "extended ($0 \leqq \text{distance} \leqq \infty$) metric space" with the metric $d(f, g) = \sup |f-g|$ then the completion of $C^\infty(R)$ via Cauchy sequences may be identified with the space $C(R)$ of all real continuous functions on R.

5. *Panel Discussion: A general curriculum in mathematics for colleges (1965 CUPM report).* Marion Cornwall, Marshalltown Community College; J. C. Mathews, Iowa State University; E. R. Mullins, Jr., Grinnell College.

6. *Cell pairs of codimension* 2, by T. M. Price, University of Iowa, Iowa City.

7. *Chains of different dimensional topologies*, by B. A. Anderson, University of Iowa.

If X is a set with cardinal at least c, there is a sequence $T_0, T_1, \cdots$ of complete metric topologies on X such that for each nonnegative integer i, T_i contains T_{i+1} and $\dim(X, T_i)=i$. Various types of these "chains" are possible. For example, if card $(X)=c$, there is a chain such that each T_i is separable; if card $(X)=c$, there is a chain such that each T_i is locally separable.

8. *Some new results in distance geometry*, by L. M. Blumenthal, University of Missouri, Columbia (by invitation).

The new results discussed are (1) the congruent imbedding of every metric $(n+2)$-tuple in the Minkowski space M_n^∞, (2) sine laws for simplices, and (3) the homothetic imbedding of all metric triples in a metric space consisting of a metric line G and a point p, not on G.

E. L. CANFIELD, *Secretary-Treasurer*

APPLICATIONS OF UNDERGRADUATE MATHEMATICS IN ENGINEERING

Early in 1967, MAA will publish a 400-page book entitled "Applications of Undergraduate Mathematics in Engineering" written and edited by Professor Ben Noble of the Mathematics Research Center, U. S. Army, University of Wisconsin. Each individual member of the Association may purchase one copy for $4.50 by sending order with remittance to the Buffalo office. Additional copies and copies for nonmembers may be purchased at $9.00 from Macmillan. The publication of this book is the direct result of activities of a panel of CUPM.

The Panel on Mathematics for the Physical Sciences and Engineering of CUPM has devoted considerable time and effort in recent years to study ways of improving the mathematics programs offered to engineering students. In examining existing courses, the committee found a perennial lack of really good engineering problems for motivation and application of mathematics. Many of the problems that appear in calculus texts, for example, although they attempt to motivate the mathematical ideas, are in reality rather superficial. They suffer by being either so overly simplified or so artificially constructed as to be meaningless from the point of view of the application.

In 1963, the Panel decided that a significant contribution could be made by making available a collection of mathematical problems drawn from real engineering situations. With the cooperation of the Commission on Engineering Education and its Executive Director, Professor Newman Hall, a large number of problems were collected from a variety of sources. These were first sorted and collected in a preliminary way by a joint committee of mathematicians and engineers. The major task, however, was the very difficult one of writing up the problems, filling in numerous gaps, and preparing a coherent exposition. It is the very good fortune of the Association that Professor Ben Noble brought his broad knowledge in both mathematics and its many fields of engineering application to the realization of "Applications of Undergraduate Mathematics in Engineering."

In the pattern of presentation adapted for this volume, each problem begins by developing the understanding of the engineering situation essentially from first principles. An approximate mathematical model is then established and carefully developed. Next, the resulting mathematical insights are applied and interpreted in the context of the original engineering situation. Numerical examples, pictorial interpretations, and generalizations of both the mathematics and the engineering are frequently possible.

The book is intended for at least three major audiences. First, for each of a number of undergraduate mathematics courses, it provides the mathematics instructor with a set of motivating problems for possible classroom presentation. Second, it brings to the attention of engineers the utility of many branches of mathematics in attacking engineering problems in their own fields. Third, the engineering student will find in this book the

opportunity to synthesize and enrich personally his educational experiences in both mathematics and engineering courses. An abbreviated version of the Table of Contents follows:

ACKNOWLEDGMENT

The Editorial Board acknowledges with thanks the services of the following mathematicians, not members of the Board, who have kindly assisted by evaluating papers submitted for publication in the MONTHLY.

T. M. Apostol, B. H. Arnold, M. G. Arsove, F. V. Atkinson, J. Auslander, R. G. Ayoub, R. W. Bagley, W. R. Ballard, P. J. Bateman, A. F. Beardon, E. F. Beckenbach, H. E. Bell, S. K. Berberian, C. Berger, R. H. Bing, T. A. Botts, D. G. Bourgin, Felix Browder, A. B. Brown, J. L. Brown, Richard T. Bumby, C. E. Burgess, R. G. Buschman, A. T. Butson, S. S. Cairns, J. C. Cantrell, L. Cesari, Chon-Yum Chao, S. Chowla, C. K. Chu, L. W. Cohen, Harvey Cohn, H. E. Conner, E. Correl, W. J. R. Crosby, C. W. Curtis, D. A. Darling, Ervin Deal, J. C. E. Dekker, R. P. Dilworth, M. D. Donsker, R. G. Douglas, Avron Douglis, Edwin Duda, L. K. Durst, H. Egan, Bernard Epstein, Trevor Evans, Carl Faith, Ky Fan, Walter Feit, Robert Finn, J. S. Frame, J. E. Freund, P. J. Freyd, M. Gardner, P. K. Ghosh, E. N. Gilbert, Leonard Gillman, Casper Goffman, R. R. Goldberg, J. K. Goldhaber, Malcolm Goldman, S. W. Golomb, H. W. Gould, D. Greenspan, R. E. Greenwood, J. H. Griesmer, H. W. Guggenheimer, S. L. Gulden, Theodore Hailperin, Alvin Hausner, Melvin Hausner, D. R. Hayes, Stevens Heckscher, Nickolas Heerema, M. R. Hestenes, E. Hewitt, Warren Hirsch, I. I. Hirschman, Harry Hochstadt, A. J. Hoffman, J. M. Horvath, A. S. Householder, S. T. Hu, B. E. Hubbard, T. E. Hull, S. Hurwitz, Nathan Jacobson, R. C. James, B. W. Jones, F. B. Jones, C. Karp, N. D. Kazarinoff, H. B. Keller, L. M. Kelly, R. B. Kelman, J. E. Kimber, W. E. Kirwan, M. S. Klamkin, V. L. Klee, A. Kleppner, H. L. Krall, S. Kuroda, J. Lambek, P. D. Lax, Solomon Leader, D. H. Lehmer, Emma Lehmer, J. Lehner, W. Leighton, D. C. Lewis, Y-F Lin, J. T. Lloyd, George Logemann, Edith H. Luchins, Y. L. Luke, G. T. McAllister, M. H. McAndrew, L. F. McAuley, P. J. McCarthy, N. H. McCoy, J. E. McLaughlin, E. J. McShane, A. G. Mackie, H. B. Mann, H. H. Martens, J. A. Mauldon, L. D. Meeker, Elliott Mendelson, D. M. Merriell, P. R. Meyer, F. Mosteller, Ronald Mullin, D. C. Murdoch, F. D. Murnaghan, Isaac Namioka, O. Ore, Barbara L. Osofsky, A. D. Otto, T. K. Pan, L. A. Pars, E. Parzen, M. H. Pearl, R. Pollack, B. Pollak, M. H. Protter, Hans Rademacher, J. F. Randolph, G. N. Raney, G. Rayna, D. K. Ray-Chaudhuri, M. O. Reade, W. T. Reid, E. Reiss, J. R. Rice, R. D. Richtmyer, John Riordan, R. M. Robinson, H. L. Rolf, P. G. Rooney, Louise J. Rosenbaum, A. Rosenfeld, N. S. Rosenfeld, Gian-Carlo Rota, N. C. Royster, Walter Rudin, Hans Sagan, W. G. Saunders, Martin Schechter, Murray Schechter, F. J. Scheid, Peter Scherk, I. J. Schoenberg, J. L. Selfridge, D. Shanks, H. N. Shapiro, R. J. Silverman, Hermann Simon, M. F. Smiley, M. R. Spiegel, R. S. Spira, E. M. Stein, K. L. Stellmacher, B. M. Stewart, F. M. Stewart, R. R. Stoll, A. S. Strauss, K. R. Stromberg, M. V. Subbarao, E. J. Taft, L. F. Takacs, Alfred Tarski, A. W. Tucker, H. G. Tucker, Atwell Turquette, W. T. Tutte, W. R. Utz, A. M. Vaidya, A. Vandeghen, J. E. Walsh, J. L. Walsh, R. J. Warne, H. F. Weinberger, G. T. Whyburn, D. V. Widder, H. S. Wilf, I. Wladaver, M. J. Wonenburger, G. S. Young, Jr., A. Zame, M. Zedek.

CALENDAR OF FUTURE MEETINGS

Fiftieth Annual Meeting, Houston, Texas, January 26–28, 1967.

Forty-eighth Summer Meeting, University of Toronto, Toronto, Ontario, Canada, August 28–30, 1967.

ALLEGHENY MOUNTAIN, West Virginia University, Morgantown, West Virginia, May 6, 1967.

ILLINOIS, University of Illinois, Urbana, May 14–15, 1967.

INDIANA

IOWA, Drake University, Des Moines, April 21, 1967.

KANSAS, Fort Hays State College, Hays, April 22, 1967.

KENTUCKY, Murray State University, Murray, Spring 1967.

LOUISIANA-MISSISSIPPI, Jung Hotel, New Orleans, Louisiana, March 4–5, 1967.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA

METROPOLITAN NEW YORK

MICHIGAN

MINNESOTA

MISSOURI, Northeast Missouri State Teachers College, Kirksville, April 29, 1967.

NEBRASKA, University of South Dakota, Vermillion, May 6, 1967.

NEW JERSEY

NORTHEASTERN

NORTHERN CALIFORNIA, University of California, Davis, February 4, 1967.

OHIO, Ohio State University, Columbus, April 22, 1967.

OKLAHOMA-ARKANSAS, Northeastern State College, Tahlequah, Oklahoma, March–April, 1967.

PACIFIC NORTHWEST, University of Montana, Missoula, June 16–17, 1967.

PHILADELPHIA

ROCKY MOUNTAIN, Western State College of Colorado, Gunnison, May 1967.

SOUTHEASTERN, Florida Presbyterian College, St. Petersburg, Florida, March 31–April 1, 1967.

SOUTHERN CALIFORNIA, San Diego State College, San Diego, March 11, 1967.

SOUTHWESTERN, University of Arizona, Tucson, March 31–April 1, 1967.

TEXAS, Austin College, Sherman, April 14–15, 1967.

UPPER NEW YORK STATE, State University College, Plattsburgh, May 20, 1967.

WISCONSIN, St. Norbert College, DePere, May 6, 1967.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, New York, N. Y., December 26–31, 1967.

AMERICAN MATHEMATICAL SOCIETY, Houston, Texas, January 24–27, 1967.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, Michigan State University, June 19–23, 1967.

ASSOCIATION FOR COMPUTING MACHINERY, Sheraton-Park, Washington, D. C., August 29–31, 1967.

ASSOCIATION FOR SYMBOLIC LOGIC, Houston, Texas, January 23–24, 1967.

CENTRAL ASSOCIATION OF SCIENCE AND MATHEMATICS TEACHERS, Chicago, November 23–25, 1967.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Houston, Texas, January 28, 1967.

OPERATIONS RESEARCH SOCIETY OF AMERICA, New York Hilton Hotel, May 31–June 2, 1967.